UiT Norges arktiske universitet

Det juridiske fakultet

Utfordringer ved bruk av kunstig intelligens i behandlingen av bevis i straffesaker

En analyse av dagens bevissystem i straffeprosessen

Bendik Wollmann

Masteroppgave i rettsvitenskap JUR-3902 juni 2020

Innholdsfortegnelse

1 Innledning

1.1 Tema

EU kommisjonens white paper «On Artificial Intelligence» av 19.02.2020: «*there is a need to examine whether current legislation is able to address the risks of AI and can be effectively enforced, whether adaptations of the legislation are needed, or whether new legislation is needed*.»[1]

Avhandlingen er en analyse av dagens strafferettslige system vedrørende bevis og bevisføring. I den forbindelse vil det redegjøres for de særskilte problemstillingene som kan følge bevis som er produsert av dataetterforskningsverktøy som benytter kunstig intelligens, herunder spesielle maskinlæringsteknikker som det vil redegjøres for i kapittel 2.

Oppgaven har to deler. Oppgavens første del vil vurdere hvordan bruken av dataetterforskningsverktøy som bruker kunstig intelligens står seg mot gjeldende regler og prinsipper i straffeprosessen. Det vil pekes på eventuelle styrker og svakheter ved det gjeldende straffeprosessuelle systemet. Oppgavens andre del vil være en de lege ferenda drøftelse som vurderer hvilke tiltak som kan bidra til å møte utfordringene kunstig intelligens fører med seg.

Avhandlingens tema er valgt på grunn av problemstillingenes nyere dato og manglende dypdykk. Norge har også blitt oppfordret fra internasjonalt hold til å utrede om teknologiens særegenhet vil få innvirkning på landenes gjeldende regelverk.[2] Det er i norske forskningsmiljøer og i «Regjeringens Strategi for kunstig intelligens» tatt til orde for at bevisbehandling gjort av datasystemer som benytter kunstig intelligens kan skape utfordringer i fremtiden.[3] Særlig er det uttrykt bekymring for manglende etterprøvbarhet, varierende kvalitet på spor og skadevirkninger dersom beslutninger tas på bakgrunn av usikkerhet og antagelser.

---

[1] EU Commission. (19.02.20) COM (2020) 65 final, *White Paper on Artificial Intelligence: a European approach to excellence and trust*, Brussel. Side 10. Sitatet er generelt formulert så selv om rapporten ikke omfatter justissektoren, har det betydning også for bruk av KI i behandlingen av straffesaker.

[2] Ibid.

[3] Norske forskningsmiljøer: Intervju med Katrin Franke, professor ved NTNU, gjort av politiforum.no. (Sist sett 21.05.2020) og Katrin Franke i og André Årsnes i Årnes, A. mfl. 1. utg. 2018. *Digital Forensics*. Hoboken: John Wiley & Sons Ltd. Side 336.
Regjeringens strategi for KI: Regjeringen. (14.01.2020), Publikasjonskode: H-2458 B, *Nasjonal strategi for kunstig intelligens*, Kommunal- og moderniseringsdepartementet, Oslo. Side 64, Tilgjengelig på: regjeringen.no.

Avhandlingen skrives derfor ut fra en hypotese om at dagens straffeprosessuelle modell og dens bevisregler kan komme til kort i møte med utfordringene kunstig intelligens fører med seg.

Det er sannsynlig at dagens modell med «fri bevisføring» og de «rettsikkerhetsmekanismene» som fungerer på en «sak til sak» basis, ikke gir adekvat dekning i møte med problemene som kan oppstå. Det tas derfor til orde i avhandlingen for at utfordringene må løses på et mer fundamentalt plan enn det dagens modell legger opp til. Oppgavens del 2 skal drøfte hvilke tiltak som vil kunne veie opp for utfordringene. Ordet «kunstig intelligens» vil noen steder forkortes til «KI» for enkelhetsskyld.

1.2 Begreper

Pålitelighet: Tradisjonelt har det i bevisteori blitt operert med begrepene pålitelighet og troverdighet for å beskrive beviset og subjektet som formidler det.[4] Påliteligheten har tradisjonelt siktet til sannhetsverdien i det som formidles, mens troverdigheten har siktet til formidleren selv.[5] Sagt med andre ord har man med pålitelighet siktet til logos, mens troverdighet har siktet til etos. Når formidleren derimot er en algoritme (f.eks. kunstig intelligens) og ikke et menneske, blir skillet noe uklart. Ved å benytte en direkte analogi ender man opp med algoritmens «out-put» [6] som det budskap det kan knyttes pålitelighet til, mens algoritmen i seg selv beskrives gjennom troverdighetsbegrepet. I norske ordbøker knyttes det imidlertid en rekke synonymer til ordet troverdighet som neppe er forenlig med en algoritme. Ord som lojal, lovlydig, sannferdig, oppriktig, rettskaffen, etc.[7] Ettersom algoritmer – i kraft å være døde[8] – ikke har fri vilje, mister ordet troverdighet mye av dens kjente karakteristikk; det subjektive akseptert er ikke tilstede. Det velges derfor – i oppgavens øyemed – å ses bort fra ordet troverdighet i beskrivelsen av «formidleren»[9] av beviset. For å dekke inn i det manglende tomrommet troverdighetsbegrepet etterlater vil ordet pålitelighet brukes i utvidet forstand. Ordet pålitelighet er i større grad en objektiv størrelse som passer bedre i beskrivelsen av en algoritme. Når det senere i oppgaven refereres til systemer som baserer seg på kunstig intelligens og de beskrives som «pålitelige», siktes det ikke bare til påliteligheten

[4] Annika Melinder på side 485 i Hedlund, M. Mfl. 1. utg. 2015. *Bevis i straffesaker: utvalgte emner.* Oslo: Gyldendal juridisk.
[5] Ibid.
[6] Med «out-put» menes i denne oppgaven det resultat som presenteres etter bruken av et digitalt program. «In-put» vil motsetningsvis være den informasjonen som går inn i det digitale verktøyet før materien analyseres.
[7] Bevis i straffesaker: utvalgte emner. Side 485.
[8] Definert som ikke levende.
[9] I denne oppgaven KI-systemet.

av det konkrete «out-put», men også programmets overordnede evne til å levere korrekte resultater. Det vil i kapittel 2 redegjøres for hvorfor en slik terminlogi er nødvendig.

Notoritet: Ordet betyr i alminnelig juridisk sammenheng etterviselig eller kontrollerbar.[10] En avgjørelse om ransaking har for eksempel notoritet når fremgangsmåten og begrunnelsen er dokumentert. I ettertid kan det føres kontroll med hvorfor avgjørelsen ble tatt og om fremgangsmåten ble gjort i henhold til gjeldende regelverk. For bevis i straffeprosessen har notoritetshensynet fått sterk rettslig stilling. Det stilles ofte strenge krav til etterprøvbarhet og kontroll for å kunne utelukke alternative forklaringer på hvorfor beviset peker i en bestemt retning.[11] Plausible alternative forklaringer kan lede til «rimelig tvil» og følgelig vil den tiltalte ikke kunne dømmes. Når det refereres til notoritet i avhandlingen siktes det til muligheten for etterprøvbarhet av det «out-put» KI-systemet har presentert. Notoritet av KI-systemets «out-put» er i den forbindelse sterkt linket til to fundamentale prinsipper i dataetterforskning som André Årnes omtaler som «evidence integrity» og «chain of custody».[12] Bevisets «integrity» sikter til det overordnede målet at beviset skal bevare sin integritet slik at bevisverdien ikke svekkes. «Forvaringskjede» (min oversettelse av «chain of custody») viser til dokumentasjonen av at bevisinnsamling, kontroll, analyse og disponering av elektroniske og fysiske beviser ivaretar det første prinsippet.[13]

Kunstig intelligens: Som det påpekes i «Når juss møter AI» finnes det ingen generelt akseptert definisjon av «kunstig intelligens».[14] Dette ikke er overraskende, all den tid det heller ikke finnes noen definisjon av organisk (eller menneskelig) intelligens, som ofte brukes som sammenligningsgrunnlag.»[15]

En av de mest anerkjent definisjonene for kunstig intelligens, også kjent som «artificial intelligence» eller AI, kommer fra Stuart Russel og Peter Norvig: *"the designing and building of intelligent agents that receive precepts from the environment and take actions that affect*

---

[10] NOU 2016: 24 Ny straffeprosesslov. Punkt. 14.2.4.3. om notoritet. På side 306.
[11] Inger Marie Sunde på side 600 i Hedlund, M. mfl. *Bevis i straffesaker: utvalgte emner.*
[12] André Årsnes på side 54 i Årnes, A. mfl. *Digital Forensics*.
[13] ibid
[14] Artikkel publisert av Bertrand Meyer, skrevet av John McCarthy, "*Communications of the ACM"* (28. Oktober 2011): https://cacm.acm.org/blogs/blog-cacm/138907-john-mccarthy/fulltext - sitert fra innledningskapitlet til House of Lords´ rapport om kunstig intelligens i Storbritannia, House of Lords (2018), «*AI in the UK: ready, willing and able?»* London. Side 8. Tilgjengelig på https://publications.parliament.uk/

[15] Bendiksen, C. Og Norman Hansen, E. 1. utg. 2019. *Når juss møter AI.* Oslo: Gyldendal Side 11.

that environment."[16] Dette betyr at programmene mottar informasjon fra et miljø og tar valg som påvirker miljøet. Ordlyden «environment» (miljø) tolkes ikke strengt og miljøet kan være et digitalt miljø. Definisjonen er generell og omfatter blant annet etterforskningssystemer som har vært i bruk relativt lenge ut ifra et teknologiperspektiv. Denne oppgaven fokuserer hovedsakelig på problemstillinger som knytter seg til digitale etterforskningsverktøy av nyere dato og da særlig de som reiser nye bevisrettslige spørsmål, jf. kapittel 2. Blant disse er systemer under grenen «maskinlæring» og da særlig de maskinlæringsmodellene som utfordrer notoritetshensynet.

Maskinlæring: Datatilsynet har definert maskinlæring i sin rapport «kunstig intelligens og personvern» slik: «*et sett teknikker og verktøy som lar maskiner «tenke» ved å lage matematiske algoritmer basert på akkumulert data».*[17]

Tom M. Mitchells definisjon av maskinlæring er en av de mest siterte og definerer maskinlæring slik: *«A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E.»*[18] Dette betyr at programmet lærer seg selv å utføre en oppgave over tid og justerer sin adferd – til det bedre – etter den forhåndsdefinerte oppgaven.

1.3 Aktualitet

I et økende antall tilfeller føres det bevis for norske domstoler som krever spesialkompetanse for å kunne forstås fullt ut, herunder digitale bevis.[19] Et av feltene som krever spesialkompetanse gjelder forståelsen av bevis som er behandlet av digitale etterforskningssystemer[20] basert på kunstig intelligens (KI-systemer), der mennesker settes på sidelinjen i kritiske faser av bevisbehandlingen.

---

[16] Russel, S og Norvig, P. 3. utg. 2010. *Artificial Intelligence: A Modern Approach*. New Jersey: Prentice Hall. Side 2. «Four possible goals: "Systems that think like humans, systems that act like humans, systems that think rationally, systems that act rationally."

[17] Datatilsynet (2018), *Kunstig intelligens og personvern*. Oslo. Side 5. Tilgjengelig på www.datatilsynet.no

[18] Mitchell, T. 1. Utg. 1997. *Machine Learning*. McGraw-Hill Science/Engineering/Math. ISBN 978-0-07-042807-2. Side 2.

[19] Nils Erik Lie på side 63 i Hedlund, M. mfl. *Bevis i straffesaker: utvalgte emner*.

[20] Digitale etterforskningssystemer er digitale verktøy som benyttes i etterforskningens øyemed, ofte for å skaffe eller behandle rådata.

Kunstig intelligens blir i dag brukt til å utføre oppgaver som er for tidkrevende eller for komplekse for mennesker, herunder bevisbehandling. Eksempelvis kan systemene gjennomgå enorme mengder digitale data i etterforskningen av et cyberangrep, brukes i analysen av DNA-sekvenser eller gjøre bilde- og stemmegjenkjenning. I tillegg til å være tid- og kostnadsbesparende brukes de også på områder der de rent kvalitativt utkonkurrerer mennesker. Eksempler på dette kan være ved atomistiske[21] analyser av bevis, for eksempel stemmegjenkjenning.

Baksiden av medaljen er at noen av systemene bruker teknikker som, gjennom sitt design, gjør det vanskelig forklare hvorfor programmet treffer en avgjørelse. Det sies av programmene er lite «gjennomsiktige». Når programmets «out-put» senere fungerer som bevis i en straffesak vil mangelen på en forklaring av *hvordan* systemet har kommet til et gitt resultat kunne komme i konflikt med notoritetshensynet i tilknytning til beviset og bevisets pålitelighet. På den måten vil en eventuell feil eller unøyaktighet være vanskelig å oppdage. Som vi skal se eksisterer det nesten alltid en ikke ubetydelig feilmargin når slike programmer benyttes.

Det kan blant annet tenkes at et bevis er fremskaffet ved hjelp av en defekt algoritme, at det «datasett»[22] programmet er trent på har mangler eller at programvaren er brukt utenfor sitt nisjeområde. Resultatet kan bli at retten legger til grunn et feilaktig eller misvisende bevisgrunnlag. Feil i bevisbildet kan i sin tur øke risikoen for et materielt uriktig resultat.

Tendensen i norsk straffeprosess er at antallet bevis som er digitalt behandlet øker.[23] Utviklingen er ventet å fortsette ettersom andelen digital informasjon hver av oss legger igjen øker. Dette er ikke overraskende ettersom digitale hjelpemidler er blitt en del av hverdagen til folk flest. De alle fleste legger i dag igjen digitale spor enten de vil eller ikke, og det i et rikelig antall.[24] For dataetterforskere er det utfordrende å manøvrere i dette havet av digital informasjon. I den forbindelse vil kunstig intelligens bidra i etterforskningen og

---

[21] En atomistisk tilnærming til bevis kjennetegnes ved at beviset granskes isolert, altså ikke i sammenheng med andre beviser. Et KI-system som analyserer ett fingeravtrykk eller én lydfil isolert vil ha en atomistisk tilnærming til beviset. Det blir bevisbedømmerens oppgave å se bevisene i sammenheng under et bevistema. Med bevistema menes det som skal bevises. Se Løvlie, A. 1. utg. 2014. *Rettslige faktabegreper*. Oslo: Gyldendal Juridisk. Side 209.

[22] «Datasett» er i denne sammenheng dataen en maskinlæringsmodell mates med i «treningsfasen» eller den dataen maskinlæringsmodellen analyserer når maskinlæringsmodellen er klar til bruk. «Datasettet» benyttes dermed som «in-put».

[23] Tendensen er i og for seg ikke overraskende all den tid antallet digitale gjenstander øker, samtidig som de legger igjen digitale spor. På samme tid flyttes mer og mer av kommunikasjonen over i det digitale rom.

[24] Eksempelvis NRKs avsløringer mai 2020 om hvor lett det er å få kjøpt digital informasjon om norske borgeres posisjoner fra mobiltelefoner på det private markedet. Tamoco var selskapet NRK kjøpte dataene fra.

saksforberedelser. Bruk av kunstig intelligens i etterforskningen vil derfor være av stor betydning for politiets evne til å bekjempe kriminalitet og bør ønskes velkommen.

Utviklingen innen teknologifeltet «kunstig intelligens» har skutt fart de siste årene og implementeringen av teknologien i det norske samfunnet er allerede i gang, og det norske politiet er intet unntak. Teknologen er i dag blitt et satsningsområde for norsk politi.[25] Verktøyene som utvikles er i flere tilfeller helt nødvendige for å holde tritt med et dynamisk kriminalitetsbilde.[26] I Interpol/UNICRI-rapporten «AI and robotics for law enforcement» fra 2019 sies det også at kunstig intelligens og robotikk er "*a present-day reality rather than a future possibility and that technological innovation continues to accelerate rapidly",* og det legges til at teknologien allerede er integrert i politiets virksomhet.[27] Å evaluere dagens regelverk opp mot de utfordringene som følger med kunstig intelligens kan således ikke la vente på seg.

Den juridiske diskusjonen rundt KI-problematikken på bevisrettens område kan sies å henge etter, og det er avhandlingens formål å bidra til debatt rundt spørsmålene. Resultatet av at diskusjonen henger etter på bevisrettens område er at verken straffeprosesslovens regler eller politipraksis har justert seg for utfordringene.[28] De spørsmålene som måtte oppstå rundt bruken av kunstig intelligens må dermed løses på bakgrunn av rettskilder som ikke har tatt hensyn til problematikken. Som vi skal se kan dette føre til at dagens system må endres for å sikre at de digitale «KI-bevisene»[29] holder samme høye standard som det stilles til andre tekniske beviser.

1.4 Avgrensninger

Oppgaven omhandler både fysiske og digitale bevis såfremt de har vært gjenstand for en digital automatisert behandling og der bevisvurderingen skjer på bakgrunn av programmets

---

[25] Artikkel på politiforum.no *«Streng kontroll med brukerne av politiets nye analyseverktøy»* (sist lest 21.05.20) https://www.politiforum.no/artikler/streng-kontroll-med-brukerne-av-politiets-nye-analyseverktoy/404492
Artikkel på politiforum.no *«Norsk forskning på framtidens politi til topps i Interpol»* (sist lest 21.05.20) https://www.politiforum.no/artikler/norsk-forskning-pa-framtidens-politi-til-topps-i-interpol/449053
National strategi for kunstig intelligens, side 64. Rapporten er utviklet for næringslivet, men kapitlet tar opp kunstig intelligens i kriminalitetsbekjempelse. Prosjektet «Ars Forensica» retter seg mot økonomisk kriminalitet.

[26] Deloitte (2018) *Policing 4.0 Deciding the future of policing in the UK*, Deloitte LLP. Side 13. Tilgjengelig på www2.deloitte.com

[27] Interpol/UNICRI (2019), *Artificial intelligence and robotics for law enforcement.* Side 19 Tilgjengelig på www.unici.it

[28] Lov 22 mai 1981 nr. 25 om rettergangsmåten i straffesaker (Straffeprosessloven, heretter strpl.).

[29] Med «KI-bevis» menes bevis som er behandlet av et datasystem som kan defineres som «kunstig intelligens» etter oppgavens definisjon.

«out-put». Med «automatisering» menes prosessen der mennesker overlater til maskiner å gjøre rådata om til et mer forståelig bevismateriale ved hjelp av et «computer system»[30]. Oppgaven avgrenses derfor mot tilfeller der en person henter ut informasjonen manuelt. Et eksempel på «automatisering» er der et dataprogram foretar fingeravtrykkanalyser fremfor at en etterforsker foretar en manuell sammenligningsanalyse. Et annet eksempel er der en automatisert programvare henter ut digitale spor fra et datasystem, mens tradisjonelle metoder ville krevd manuell identifisering og sikring av en dataetterforsker.

Hovedtyngden av oppgaven vil ligge på KI-systemenes kapasiteter og begrensinger når bevis analyseres og funnene presenteres. I den forbindelse vil Flaglienes dataetterforskningsmodell benyttes. I Flagliens modell er «dataetterforskningsprosessen»[31] delt opp i fem steg. (1) identifisering, (2) sikring, (3) klargjøring, **(4) analyse og (5) presentasjon**.[32] Oppgavens hovedfokus vil ligge på de to siste. Merk at Flagliens modell er utviklet for digital dataetterforskning og stegene knytter seg utelukkende til digital informasjon, noe som medfører at fysiske bevis (eksempelvis et fingeravtrykk) faller utenfor. Imidlertid vil spørsmålene som reises i analyse- og presentasjonsfasene være de samme. Det er derfor hensiktsmessig å benytte modellen i denne avhandlingen.

Søkelyset rettes kun mot den straffeprosessuelle regulering av bevis som er innhentet. Rettslige spørsmål knyttet til innhenting av bevis faller således utenfor. Det samme gjør bevis som føres for retten i andre spørsmål enn straff, eksempelvis ved begjæring om kommunikasjonskontroll.

Selv om analysen skjer ut fra et straffeprosessuelt perspektiv vil svakheter ved beviset ha størst skadepotensial dersom det fører til at uskyldige dømmes. Siden kjernen i KI-problematikken langt på vei er den usikkerhet systemene opererer med vil det rettes særlig oppmerksomhet mot bevisets «robusthet».[33] Robusthetskravet er en del av det

---

[30] Definisjonen av «computer system»" er hentet fra definisjonen inntatt i Cybercrime konvensjonen art. 1 a:
«computer system" means any device or a group of interconnected or related devices, one or more of which, pursuant to a program, performs automatic processing of data;
Convention on Cybercrime, Budapest, 23.11.2001. Tilgjengelig på:
https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/185

[31] Dataetterforskningsprosessen er en oversettelse. Originalt: «the digital forenics process»
[32] Anders O. Flaglein på side 66 i Årnes, A. mfl. *Digital Forensics*.
[33] Strandberg, M, (2010) *Beviskrav i sivile saker. En bevisteoretisk studie av den norske beviskravslærens forutsetninger*. (Doktoravhandling, Universitetet i Bergen).
Strandberg definerer robusthet som «fravær av skjørhet overfor ny relevant informasjon» på side 412.

strengebeviskravet i strafferetten og er ment å ivareta tiltaltes interesser.[34] Oppgaven skrives derfor – hovedsakelig – ut ifra et rettssikkerhetsperspektiv, men analysen vil samtidig se hen til rettssystemets interesser for øvrig.

1.5 Rettskildebilde og metode

Metodemessig vil del 1 inneholde en analyse av dagens rettstilstand, de lege lata. Det vil også bli pekt på styrker og svakheter med dagens system og hvordan systemet ivaretar eller utfordrer de grunnleggende hensynene straffeprosessen bygger på. I del 2 vil det drøftes mulige endringer til dagens straffeprosessuelle system, de lege ferenda. I den forbindelse vil del 2 være en rettspolitisk øvelse med det formål å bidra i løsningen av problematikken.

Rettskildebildet som skal analyseres vil i hovedsak være straffeprosesslovens regler sett i sammenheng med Grunnlovens bestemmelser om menneskerettighetene, og Den europeiske menneskerettighetskonvensjon (EMK).[35] Oppgaven er langt på vei en analyse av hvordan dagens strafferettslige system ivaretar problematikken rundt KI-bevisets pålitelighet når programvaren, gjennom sitt design, ikke tillater den grad av notoritet vi er vant med i dagens bevisføring. I den forbindelse må de rettssikkehetsmekanismene som skal fremme et materielt riktig resultat vurderes, herunder påtalemyndighetens ansvar for et forsvarlig avgjørelsesgrunnlag, innsyn- og kontradiksjonsreglene, samt rettens ansvar for sakens opplysning. Siden spørsmålet omhandler en spesialdel innenfor data vil også sakkyndighetsordningens betydning drøftes.

2 Del 1 – Analyse av rettstilstanden – de lege lata

2.1 Hvordan kunstig intelligens kan påvirke bevis i straffesaker – de faktiske forholdene

For å undersøke hypotesen om at dagens straffeprosessuelle ordning ikke gir adekvat dekning i møte med KI-problematikken, er det nødvendig med et innblikk i hvilke utfordringer KI-teknologien fører med seg. Utfordringene må derfor identifiseres. For sammenhengens del vil det aller først bli knyttet noen korte bemerkninger til hvorfor kunstig intelligens brukes i etterforsknings øyemed.

---

[34] Robusthetskravet skiller seg fra andre utredningskrav ettersom hensynet bare tilgodeser den tiltalte. Eksempelvis vil kravet til robusthet skille seg fra rettens ansvar for sakens opplysning etter strpl. § 294 da hensynene etter § 294 i utgangspunktet ivaretar hensyn i begge retninger, ikke bare i tiltaltes.

[35] Lov 17 mai 1814 Kongeriket Norges Grunnlov. (Grunnloven, heretter Grl.)
Convention for the Protection of Human Rights and Fundamental Freedoms, Rome, 04.11.1950. (Menneskerettskonvensjonen, heretter EMK),

2.1.1 Hvorfor benytte kunstig intelligens i etterforskning

Teknologien gjør det mulig å bekjempe kriminalitet på andre måter enn tidligere. Effektiviseringshensynet blir ofte trukket fram som hovedgrunnen for å ta i bruk kunstig intelligens i politiet. Eksempelvis benytter politiet i dag skriftgjenkjenning basert på maskinlæring for å identifisere og lokalisere kjøretøyer som er i politiets søkelys ved å gjenkjenne skiltnumre.[36] Videomaterialet som analyseres har lenge vært tilgjengelig for politiet, men det har ikke vært en prioriteringssak å sette av flere hundre politibetjenter til å se gjennom det. Ved siden av effektivitetshensyn er det samtidig ønskelig å benytte KI-systemer på felter der de utkonkurrerer mennesker kvalitativt, snarere enn kvantitativt. Et eksempel kan være identifisering bakgrunnsstøy i en telefonsamtale. En analyse av støyen kan indikere hvor en person ringer fra eller hvem som snakker i bakgrunnen, ved hjelp av lyd- eller stemmegjenkjenning. Dette er oppgaver KI-systemer kan analyse bedre enn mennesker og dermed er bruken også begrunnet i rettferdighetshensyn og politiets evne kriminalitetsbekjempelse.

I andre sammenhenger har politiet ikke annet valg enn å bruke disse automatiserte systemene dersom de skal ha muligheter for å møte dagens – og fremtidens – kriminalitetsbilde. Selv om de kraftigste digitale verktøyene i skrivende stund benyttes, kan det ta flere år å gjennomgå rådataen i enkelte saker for å identifisere bevis. Eksempelvis er den største pågående saken til Økokrim nå på svimlende 52.0TB[37] med bevismateriale. Til sammenligning utgjorde Panama Papers dokumentene 2.6TB, mens 1,7 millioner e-poster utgjør omlag 160GB[38]. I Panama saken brukte 376 journalister fra 100 mediehus ett år på å gå gjennom dokumentene på 25 forskjellige språk.[39] Når saker når dette omfanget er det forståelig at mesteparten av den digitale etterforskningen må overlates til kunstig intelligens.

Riksadvokaten har uttrykt bekymring for mengden informasjon som kan inngå som saksmateriale i dag, og da særlig i større økonomisaker. Det uttales i den forbindelse at *«[m]etodeutviklingen og prioritering av ressurser fremover må innrettes mot å effektivisere*

[36] Automatisk skiltgjenkjenning (ANPR) Tilgjengelig på: https://www.vegvesen.no/fag/fokusomrader/trafikksikkerhet/skiltleser
[37] TB er forkortelsen til «terabyte». 1 terabyte tilsvarer 1 000 gigabyte.
[38] GB er forkortelsen til «gigabyte». 1 gigabyte tilsvarer 1 000 megabyte.
[39] Riksadvokatens notat 07.01.2019. *Notat om utviklingen ved etterforskningsfelt.* Punkt 4.3.1. Side 9.

mengdehåndtering for eksempel ved hjelp av kunstig intelligens (AI) og språkanalyser (NLP).»[40]

2.1.2 Utfordringer ved bruk av kunstig intelligens i bevisbehandling

Det er flere grunner til at det kan være utfordrende å benytte kunstig intelligens i bevisbehandling. I det videre skal søkelyset rettes mot noen av dem.

2.1.2.1 Koder

Alle dataprogrammer[41] er bygd opp av koder og kodene er i og for seg ikke en særegenhet med kunstig intelligens. Imidlertid følger det en mulighet for at feil i kodene kan utgjøre en feilkilde når programmet analyserer eller behandler beviset. Det er derfor ønskelig å kunne kontrollere kodene til et program for å avfeie alternative forklaringer, herunder at programmets «out-put» skyldes en feilkilde i kodene. Kodene som bygger opp programmer kalles ofte «kildekode» og kildekodene kan holdes skjult for brukerne eller være «åpne».

KI-systemer er bygd opp av koder som igjen bestemmer hvilke algoritmer[42] programmet skal benytte. Når programmet «skrives» kan potensielle feil i kodene påvirke KI-systemenes «out-put». Programmets «out-put» kan utgjøre bevismaterialet og det er dermed fare for at KI-systemet presenterer et misvisende, unøyaktig eller feilaktig bevis. Feil i programvarer blir ofte referert til som «bugs» og kan ha større eller mindre innvirkning på systemets funksjoner. En tommelfingerregel er at alle programmer inneholder «bugs» og at antallet «bugs» er mer eller mindre er proporsjonalt med systemets kompleksitet.[43] Det følger da logisk at alle programmer inneholder potensielle feilkilder som muligens kan påvirke programmets funksjoner.

Det kan argumenteres for at småfeil og ikke-optimale algoritmer ikke har noen reell betydning dersom programmet fungerer når det blir tatt i bruk. Det er samtidig bort imot umulig å ha kontroll over alle potensielle feilkilder i et komplekst dataprogram. Resultatet er at det aksepteres en viss usikkerhetsmargin når mer komplekse programmer brukes. For de mer sofistikerte maskinlæringssystemer er det vanlig å operere med en «error rate»[44] som tar

---

[40] Ibid.
[41] Med «dataprogrammer» siktes det til «software».
[42] Algoritme, i matematikk og databehandling en fullstendig og nøyaktig beskrivelse av fremgangsmåten for løsning av en beregningsoppgave eller annen oppgave. Hentet fra store norske leksikon 8. april 2020.
[43] Saltzer, J. and Frans Kaashoek. 1. utg. 2009. *Principles of Computer System Design: An Introduction.* Burlington. Morgan Kaufmann Publisher. Side 19.

[44] «Error rate» er frekvensen av falske positive og falske negative. På norsk kan «feilrate» benyttes. «Error raten» refereres ofte til i prosent.

hensyn til denne feilmarginen og kvantifiserer den. På den måten vil det være mulig å si noe om programmets generelle pålitelighet. Helt enkle programmer med få oppgaver ligger godt an til å ha en lav «error rate». Blant annet fordi de har færre linjer med koder og fordi oppgavene de skal løse ofte er relativt enkle, for eksempel speilkopiere all digital informasjon på en mobiltelefon. For å utføre mer komplekse oppgaver benyttes ofte selvlærende systemer og dermed øker antallet potensielle feilkilder.[45] Imidlertid har mer avanserte programmer vist seg veldig fordelaktige i dataetterforskning og potensialet er enormt.

2.1.2.2 Treningsdata

Systemer som baser seg på maskinlæring trenes ofte opp på store datamengder, også kalt «treningsdata». I denne treningsfasen kan rådataen programmet trener på ha mangler og manglende vil kunne inngå i den ferdige maskinlæringsmodellen.[46] Et annet problem er at treningsdataen ikke gjenspeiler det datasettet programmet senere brukes på, noe som igjen vil påvirke programmets «out-put». Treningsdataen kan dermed i seg selv utgjøre en feilkilde.

Kunstig intelligens som baserer seg på maskinlæringsmodeller som «nevrale nettverk» og «dyp læring», benytter ofte mange parametere. Et høyere antall parametere gjør det generelt vanskeligere å forklare programmets «out-put». Dette gjør det vanskeligere å finne potensielle skjevheter eller feil i treningsdataen. En annen problemstilling det er viktig å ta i betraktning er at systemene ofte er svært ømfintlige for eksternt «støy». Støy vil her si alt som programmet ikke spesifikt er trent opp på. Et lite avvik fra den spesifikke treningsdataen kan ha innvirkning på resultatet, noe eksemplene nedenfor tydeliggjør.[47] Det er med andre ord viktig å ha en forståelse av hvilke treningsdata som er benyttet slik at bruken av KI-systemet begrenses til nøyaktig samme oppgaver. Det er på denne bakgrunn rimelig å påstå at systemene – og bruken – må være gjenstand for kontroll dersom de skal produsere bevis til straffesaker på en stor skala.

Et eksempel fra Tyskland er illustrerende for hvor skjøre programmene kan være for avvik.[48] I et lab-eksperiment lærte en forskergruppe et KI-system å identifisere fingeravtrykk. Programmet så ut til å fungere med en relativt lav «error rate». Da verktøyet en dag ble testet av kinesiske utvekslingsstudenter reagerte den med å presentere langt flere falske resultater.

---

[45] Som hovedregel. Unntak kan selvsagt tenkes.
[46] Naarttijärvi, M. (2017). *Rättsstatlighet och algoritmiska svarta lådor*. I: Örjan Edström, Johan Lindholm & Ruth Mannelqvist (ed.), Jubileumsskrift till Juridiskainstitutionen 40 år (s. 245-259). Umeå: Juridiska institutionen, Umeå universitet. Side 252.
[47] I litteraturen beskrives termen som «narrow AI».
[48] Personlige meddelelser, Katrin Franke, professor i computer science ved NTNU. 2019 og 2020.

Det viste seg at treningsdataene stort sett var fingeravtrykk av sentraleuropeiske borgere som hadde små variasjoner seg imellom sammenlignet med kinesiske fingeravtrykk. Eksemplet får fram et generelt problem med sofistikerte selvlærende systemer, nemlig hvor oppgavespesifikke de er. For å bygge et robust fingeravtrykkprogram som baserer seg på selvlærende systemer kreves det at KI-systemet trenes opp på et bredt representativt utvalg av verdens befolkning, ettersom en gjerningmann kan ha hvilken som helst avstamning.

En annet eksempel gjelder en KI-programvare som var utviklet for å skanne pass.[49] Et problem som raskt ble klart var at programvaren så ut til å gi flere falske resultater etterhvert som tiden gikk. Etter en stund fant de årsaken. Lampen på pulten hadde over tid gått fra hvit til et mer gulaktig lys, noe som igjen hadde påvirket «in»- og dermed «out-put». Lyspæren måtte derfor skiftes ut jevnlig, men resultatene ble ikke helt optimale. Brukerne av programmet fikk eksempelvis forskjellige resultater med samme pass alt ettersom de skannet passet før eller etter lunsj. Etter en stund fant de årsaken; solen som skinte inn gjennom vinduene flyttet seg på himmelen etterhvert som arbeidsdagen gikk. Dette hadde igjen innvirkning på hvordan programmet tolket «in-put»-et. Eksemplet illustrerer godt hvor sensitive og «smale» («narrow») slike KI-systemer er, og hvorfor det også må føres kontroll med bruken, ikke bare med programvaren selv.

En annen utfordring er selvlærende systemers tendens til å bygge opp «bias»[50] (skjevheter) i treningsfasen og vektlegge disse i resultatet.[51] Fenomenet forekommer fordi KI-systemer ser etter korrelasjoner snarere enn kausalitet. Et konkret eksempel kommer fra USA og gjelder dataverktøyet COMPAS. Dataverktøyet baserer seg på maskinlæring og er et støtteverktøy for dommerne i amerikansk rett.[52] Programmet estimerer sannsynligheten for at den tiltalte vil begå nye lovbrudd i fremtiden, noe som har direkte innvirkning på straffeutmålingen.[53] Siden KI-systemet er trent opp på empiriske data (vandel) og spørreundersøkelser, fant dataverktøyet en korrelasjon mellom det å være afroamerikansk og å begå nye lovbrudd i

---

[49] Personlig meddelelse, Katrin Franke (2019).
[50] Begrepet algoritmisk bias beskriver systematiske og repeterbare feil som skaper urettferdige utfall. Friedman, B. og Nissenbaum, H. (1996) *Bias in computer systems*. Publisert: ACM Transactions on Information Systems, juli 1996. Side 332.
[51] Bass, D. and Huet, E. (2017). *Researchers Combat Gender and Racial Bias in Artificial Intelligence.* Tilgjengelig på https://www.bloomberg.com/news/articles/2017-12-04/researchers-combat-gender-and-racial-bias-in-artificial-intelligence (sist lest 29.05.20)
[52] Angwin, J., Larson, J., Mattu, S. & Kirchner, L. (2016) *Machine Bias.* Tilgjengelig på: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (sist lest 29.05.20)
[53] Eric Loomis vs State of Wisconsin, cert. denied, 137 S.Ct. 2290 (2017). Wisconsin Supreme Court. Appeal to the United States Supreme Court denied. Se for eksempel avsnitt 75. Tilgjengelig på: https://www.courts.ca.gov/documents/BTB24-2L-3.pdf

større grad enn for hvite. Dette omtales som «bias» ettersom hudfarge ikke er en direkte årsak til kriminalitet, da det er andre mekanismer som gjør at afroamerikanere er mer involvert i gjengangerkriminalitet. Eksempelvis vil diskriminering i samfunnet føre til at afroamerikanere generelt sett lever med en strammere økonomi enn hvite. Det strider derfor mot rettsfølelsen å legge vekt på hudfarge, kjønn, religion, etc. Et verktøy som dette benyttes per i dag ikke i Norge, men utfordringene knyttet til kvaliteten på treningsdata og hva KI-systemer trekker ut av dataen, er universelle utfordringer når sofistikerte maskinlæringsprogrammer trenes. Å ikke føre adekvat kontroll med hvilke treningsdata som benyttes vil derfor kunne føre til uønskede og ubegrunnede skjevheter i programmets «out-put».

2.1.2.3 «Svart boks»-problematikken

Det kan dermed slås fast at det er nødvendig å føre kontroll med slike selvlærende programmer grunnet deres mulige feilkilder. Eksempelvis feil i kodene, skjevheter i treningsdataen og det svært begrensede bruksområdet. Direkte kontroll med programmet og dens avgjørelser kan imidlertid møte utfordringer gjennom «svart boks»-problematikken.

«Svart boks»-problematikken er et fenomen som følger selvlærende systemer som operer med et større antall parametere.[54] Problematikkens kjerne er KI-systemets manglende evne til å gi en forklaring på hvorfor programvaren treffer et gitt resultat. Problemet oppstår når en programvare overlates til å trene seg selv ved at «*algoritmen själv definierar eller modifierar de regler som gäller för beslutsfattandet.»*[55] Det store antallet datapunkter og justeringer gjør det nærmest umulig for et menneske å forstå hvordan modellene danner seg referansepunkter og gjør avveininger. Dette gjelder som nevnt særlig ved bruk av teknikker som «nevrale nettverk» og «dyp læring» der et større antall parametere og «lag»[56] er brukt. Dersom antallet parametere holdes relativt lavt kan systemene la seg forklare, men det går en grense.[57] Nøyaktig hvor denne grensen går er et spørsmål for videre forskning.

---

[54] EU Commission. (08.04.2019). *Ethics Guidelines for Trustworthy* AI, High-Level Expert Group on Artificial Intelligence. Brussels. Nederst på side 21.
[55] Naarttijärvi, M. *Rättsstatlighet och algoritmiska svarta lådor*. Side 252. Setningene mellom fotnote 37 og 38.
[56] Nevrale nettverk har forskjellige lag med datapunkter. Noen av lagene kan være skjulte og refereres da til som «hidden layers».
[57] Shalaginov, A. Og Franke, K. (2017) *A deep neuro-fuzzy method for multi-label malware classification and fuzzy rules extraction*. Publisert i: 2017 IEEE Symposium Series on Computational Intelligence (SSCI). Seksjon 2, bokstav A.

Grunnen til at flere parametere ønskes benyttet er at de mer sofistikerte modellene har en tendens til å være de mest fordelaktige. Fordelen – sammenlignet med «ikke selvlærende systemer» – ligger i systemenes evne til å finne mønstre på egenhånd, noe maskinene ofte gjør bedre en mennesker når programmeringsutfordringene blir for komplekse.[58] Dagens maskinkraft og tilgang på data gjør maskinlæringsmodellene mer aktuell enn når idéen ble utviklet på 1950-tallet. De mest sofistikerte modellene er dermed – gjennom sitt design – ofte uatskillelig fra «svart boks»-problematikken, og i skrivende stund kan en i de fleste tilfeller ikke få den ene uten den andre. I et bevisperspektiv skaper dette utfordringer for notoriteten til selve beviset, som igjen går ut over bevisets pålitelighet. For å sikre en rettferdig rettergang når det gjøres inngrep i den tiltales rettigheter, for eksempel retten til kontradiksjon, kreves det normalt adekvate «counterbalancing factors».[59] Myndighetene skal gjennom mottiltak sørge for at den tiltalte har hatt en «reell mulighet» til å forsvare sine interesser.

I forbindelse med «svart boks»-problematikken diskuteres det i litteraturen hvorvidt beslutningstaking gjort av mennesker også lider under av å være en «svart boks», der beslutningstakeren selv ikke vet nøyaktig hvorfor avgjørelsen ble tatt, men rasjonaliserer valget tilstrekkelig slik at valget lar seg forsvare. Se for eksempel Rachel A. Searston og Jason Chin «The Legal and Scientific Challenge of Black Box Expertise».[60] Analogiverdien ligger i at vi aksepterer avgjørelser truffet av eksperter fra for eksempel sakkyndige, selv om litteraturen tyder på at de heller ikke vet nøyaktig hvordan de traff avgjørelsen. En slik forståelse taler for at sammenligningsgrunnlaget vårt må være de menneskelige aktørene i straffeprosessen, snarere enn et ideelt scenario.

Hvordan «svart boks»-problemet oppstår og hva det kan bety i praksis

Kort fortalt benytter nevrale nettverk seg av – som navnet indikerer – et nett av datapunkter som er koblet sammen. Antallet datapunker kan overstige millioner. Hvert datapunkt har en verdi som knytter den sterkere eller svakere til et annet punkt i treningsfasen. Kompleksiteten blir dermed årsaken til at logikken eller resonnementet til maskinen ikke lar seg forklare

---

[58] Å programmere inn komplekse mønstre er vanskelig. Å la en maskin gjenkjenne mønstre på egenhånd, gjennom eksempelvis 100 000 eksempler, gjør at programmet selv definerer reglene for beslutningsprosessen. Skaperen av algoritmen gir da i fra seg kontrollen og programmet bestemmelser selv hvilken logikk den skal bruke, jf. Naarttijärvi, M. Rättsstatlighet och algoritmiska svarta lådor. Side 252.

[59] Se for eksempel EMDs tankesettet i Al Khawaja and Tahery v United Kingdom avsnitt 113-118.
For EMDs forståelse av hvilke "counterbalancing factors» som inngår i den tiltaltes rett ved søkeverktøy i «stordata», se Sigurður Einarsson and others v. Iceland, 4 September 2019, (Saksnummer 39757/15). Avsnitt 85-92.

[60] Searston, R. A. og Chin, J. (2019) *The Legal and Scientific Challenge of Black Box Expertise*. Publisert i: University of Queensland Law Journal. The University of Sydney Law School.

(verken av KI-systemet selv eller av mennesket som bruker det). Det en står igjen med er hva som går inn i programvaren og hva som kommer ut – selve mekanismen er skjult og utilgjengelig – dermed navnet «black box» eller «svart boks».

En analogi til verden rundt oss er dersom man aldri har benyttet et dørhåndtak før. Man lærer fort hva som skal til for å åpne døren, man presenterer et «in-put» (man dytter ned håndtaket og dytter) og blir presentert for resultatet eller «out-put»-et (døren åpner seg). Mekanismen inni døren er man like uvitende om, og vi trenger heller ikke sette oss inn i hvordan mekanismen fungerer så lenge vi erfarer at den fungerer. Det er langt på vei slik vi benytter oss av de KI-systemene som er «svarte bokser». Så lenge programmet gjør det den er programmert til (f.eks. åpne dører) med en akseptabel «suksess rate»[61], er programmet nyttig for oss. Når det gjelder de mest sofistikerte maskinlæringsmodellene vil som sagt modellens resonnement være bortgjemt i det store antallet datapunkter. En manglende rasjonell forklaring til hvorfor et bevis peker i en gitt retning vil i utgangspunktet kunne komme i konflikt med det strenge beviskravet i strafferetten.

Mangelen på notoritet kan også være en alvorlig brist i adgangen til å imøtegå beviset, noe som kan være nødvendig ettersom andelen «falske positive» er relativt høyt.[62] Det sier seg selv at det i utgangspunktet er vanskelig å bevise at det er gjort en feil når feilen ikke er åpenbar, og begrunnelsen ikke er tilgjengelig.[63] Antallet potensielle feilkilder gjør det viktig at man ikke neglisjerer utfordringene. Det må derfor føres en strengere kontroll med KI-systemer som er såkalte «svarte bokser» enn andre dataetterforskningsverktøy. Dersom utfordringene neglisjeres ved at man for eksempel fortsetter å behandle KI-programmene på samme måte som tradisjonelle IT-verktøy, står man over for potensielt store rettssikkerhetsproblemer.[64]

---

[61] Motsetningen til «error rate». Ratene er omvendt proporsjonale.
[62] Huber, J-A., Memminger, M., Soppitt, M. og Hayday, M. (2019). *Cutting Through Complexity In Financial Crimes Compliance*. Se innledende del av artikkelen: Tilgjengelig på: https://www.forbes.com/sites/baininsights/2018/02/14/cutting-through-complexity-in-financial-crimes-compliance/#17193147588d (Sist lest 21.05.20)
[63] Åpenbare feil vil umiddelbart bli avfeid og det er lite sannsynlig at slike feil noen ganger når en hovedforhandling, og langt mindre lagt til grunn. Et eksempel kan være at en laser som måler hastigheten til kjøretøy presenterer et «out-put» som tilsier at det målte kjøretøyet holdt en hastighet på 800 km/t.
[64] Riksadvokaten har tatt opp utfordringer knyttet til rettssikkerheten i teledataskandalen og i Crash Cube (vegvesenet) saken. Se: Riksadvokaten (20.12.19) *Statens vegvesen og deres bruk av dataverktøy ved sakkyndig bistand i straffesaker.* Riksadvokaten (20.12.19) *Statens vegvesen og deres bruk av dataverktøyet CrashCube.* Riksadvokaten (15.01.20) *Trafikkdata innhentet fra Danmark - videre oppfølging*. Side 2. «Riksadvokaten vil likevel understreke viktigheten av fortsatt bevissthet om de mulige feilkilder som kan foreligge både ved innhentingen, bearbeidingen og bruken av slik informasjon.»

Ved å akseptere tilstedeværelsen av «svart boks»-problemet, vil man samtidig akseptere at det enkelte resultat («out-put») ikke kan begrunnes konkret. Det nest beste vil i så måte være å si noe om programmets treffsikkerhet på generelt grunnlag. En slik sannsynlighetstilnærming gjør at man indirekte kan si noe om påliteligheten av det enkelte «out-put». Det er på samme tid avgjørende at denne informasjonen faktisk når frem til aktørene i straffeprosessen slik at de kan vurdere KI-bevisets beviskraft i lys av den generelle påliteligheten systemet operer med.

Et eksempel på en strafferettslig prosess der KI-bevis føres kan være: Påtalemyndigheten fører et bevis der et stemmegjenkjennende KI-system med base i «dyp læring» (en maskinlæringsteknikk med mange parametere), har gitt et positivt treff på tiltaltes stemme. Lydopptaket KI-systemet har analysert er en telefonsamtale mellom to personer og programvaren mener tiltaltes stemme er å høre i bakgrunnen av samtalen. Programvarens treff indikerer at tiltalte var rundt telefonen på gjeldende tidspunkt. Telefonens posisjon kan lokaliseres til gjeldende «åsted», og bevistemaet er hvorvidt tiltalte var på åstedet på gjerningstidspunktet. Dersom lydopptaket er av en slik karakter at dommerne ikke klarer å identifisere tiltales stemme blir spørsmålet hvor pålitelig programmet er. Programmet som er trent opp på menneskestemmer klarer imidlertid ikke å forklare hvorfor den har fått et positivt treff. Dersom det kommer en innsigelse fra forsvarer om at treffet må være en «falsk positiv», må det vurderes. Bakteppet for vurderingen er at enhver «rimelig tvil» skal komme tiltalte til gode. Spørsmålet blir da hvordan det skal avgjøres hvorvidt slike programmer er pålitelige.

I eksemplet må det vurderes hvor pålitelig KI-systemets «out-put» må antas å være. En treffprosent på 85 % (15 % «error rate») vil eksempelvis ikke være tilstrekkelig for oppfyllelse av beviskravet, dersom lydfilen er det eneste fellende beviset i saken, jf. at enhver rimelig tvil om faktiske forhold skal komme tiltalte til gode.

I en bevistvist som den ovenfor blir KI-systemet i seg selv tvistegjenstand sammen med dets «out-put». Selv om slike programmer generelt sett gjør en bedre jobb enn rettens aktører i å identifisere stemmer, betyr det ikke at programmet er feilfritt. Jonas Ekfeld har i sin doktoravhandling «om informationtekniskt bevis» vist at det er utfordringer knyttet til en rekke feilkilder på informasjonsteknologien område.[65] Alle hans funn skal ikke gjengis her, men når Ekfeld treffer den konklusjon at antallet feilkilder er betydelig i tradisjonell

[65] Ekfeldt, J. (2016). *Om informationstekniskt bevis.* (Doktoravhandling, Stockholms universitet). Punkt 7.2, side 449 – 455. Tilgjengelig på: http://su.diva-portal.org/smash/get/diva2:900594/FULLTEXT05.pdf

dataetterforskning, vil det neppe være urimelig å påstå at problemene ikke vil forsvinne når systemene automatiseres, snarere tvert imot.[66] Samtidig er bruken av KI-systemer er ventet å øke, jf. Riksadvokatens uttalelse om at *«[m]etodeutviklingen og prioritering av ressurser fremover må innrettes mot å effektivisere mengdehåndtering for eksempel ved hjelp av kunstig intelligens.»*[67] Det vil på bakgrunn av en slik forståelse være avgjørende å kunne føre kontroll med KI-systemene som brukes.

2.1.3 Vurdering av systemenes pålitelighet

Fra et teknisk perspektiv er det hovedsakelig tre måter å undersøke en «svart boks»-KI og dermed si noe om programmets pålitelighet. Kun en av måtene gir oss en «error rate», mens de to andre handler om å påvise feil i selve systemets oppbygning. Dersom det kan påvises feil som kan ha virket inn på programmets «out-put» vil påliteligheten til programmet – og dets «out-put» – falle drastisk. Det er imidlertid ingen automatikk i at et program som inneholder feilkilder ikke kan levere et nøyaktig bevis, men påtalemyndighetens oppgave vil på en annen side bli mye vanskeligere grunnet det strenge beviskravet, jf. at rimelig tvil skal komme den tiltalte til gode.

For «svart boks»-systemer vil det enkleste være å se gjennom programmets treningsdata for å se om de inneholder åpenbare svakheter, jf. fingeravtrykkeksemplet ovenfor. Dersom en ikke kan avdekke svakheter i treningsdataene vil det neste være å teste programmet. Dette innebærer at testerne får tilgang til programmet og kan teste det på uavhengige datasett innenfor det domene programmet skal brukes. Det tredje og mest tidkrevende vil være å få tilgang til programmets koder og gjennomgå disse.

Samtidig som det kreves tilstrekkelig kompetanse for å gjennomgå avanserte dataetterforskningsprogrammer vil det også være tidkrevende, og ofte mer tidkrevende enn å programmere selve programmet. At automatisering fører med seg mye tids- og ressurskrevende etterarbeid er kjent som «the paradox of automation». En avhandling fra NTNU med tittelen «The Paradox of Automation in Digital Forensics» forklarer fenomenet

---

[66] White Paper 19.02.20. *On Artificial Intelligence - A European approach to excellence and trust*. Side 12: «As with the risks to fundamental rights, these risks can be caused by flaws in the design of the AI technology, be related to problems with the availability and quality of data or to other problems stemming from machine learning. While some of these risks are not limited to products and services that rely on AI, the use of AI may increase or aggravate the risks.»

[67] Riksadvokatens notat 07.01.2019. *Notat om utviklingen ved etterforskningsfelt.* Punkt 4.3.1. Side 9.

slik: *«the more efficient the automated system, the more crucial the human contribution of the operators. Humans are less involved, but their involvement becomes more critical".*[68]

Merk at et dataetterforskningsprogram er anbefalt å testes gjennom hele sin levetid da hver oppdatering i prinsippet kan føre med seg potensielle feilkilder.[69] Det må derfor settes av mye ressurser for å kontrollere programmene man benytter, enten de er innkjøpte eller selvutviklede. Fra et ressursperspektiv vil det at programmene fungerer gjerne være grunn nok for å bruke dem. Når bevisene programmene tilbyr heller ikke blir motsagt i utstrakt grad vil interessen av å iverksette tid- og ressurskrevende falsifiseringsrutiner være av liten. At prosedyrene er tid- og ressurskrevende betyr imidlertid ikke at de ikke er nødvendige.

Programmene testes åpenbart hos leverandøren, men det opplyses ikke overfor kjøperne om hvor treffsikkert programmet er. At programmene er gode nok til å selges, betyr altså ikke at de ikke inneholder feil. Dersom den underliggende treffsikkerheten i programmet er 99 % vil programmet i løpet av 1000 tester i snitt gi 10 falske resultater. 99 % regnes som en svært høy «suksessrate» innenfor kunstig intelligens og en må regne med at programmene ikke er fullt så treffsikre, jf. eksemplene ovenfor. Dette støttes av Graeme Horsmans artikkel fra 2018.[70] I hans undersøkelser gjort på dataetterforskere uttales det på bakgrunn av spørsmål nr. 10 i undersøkelsen: «*Q10 indicates that on average, respondents suggested that an 'error rate' of approximately 10% is acceptable. This can be viewed as both a realistic target and also a concerning factor that practitioners may be willing to accept such levels of inaccuracy in tools designed to support the establishment of fact.*» Den «suksessraten» som de fleste av deltakerne anser som akseptabel og realistisk faller altså på 90 %. Det er på det rene at usikkerheten ikke er så lav som en skulle ønske og at det i utgangspunktet kan medføre vansker når beviskravet skal oppfylles.

Resultatet av eventuelle undersøkelser av koder, treningsdata og testing vil kunne si noe om programmet inneholder åpenbare skjevheter, hvilken «error rate» et program opererer med og andre potensielle mangler. En vil dermed kunne skape seg et bilde av programmet «generelle pålitelighet», som kan fungere som supplement for den manglende notoriteten til det enkelte

---

[68] Borhaug, T. S. (2019). *The Paradox of Automation in Digital Forensics.* (Masteroppgave, NTNU, Norges teknisk-naturvitenskapelige universitet). Side 22. Tilgjengelig på: https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2617753 Definisjonen er utarbeidet av Josh Kaufman.
[69] EU kommisjonens White Paper, *on Artificial Intelligence: a European approach to excellence and trust*. Side 20.
[70] Horsman, G. (2019) *Tool testing and reliability issues in the field of digital forensics.* Publisert i: Digital Investigation 01.03.2019. Side. 163-175.

bevis. Notoritetshensynet søkes således å ivaretas indirekte. Det er tatt til orde for at dataetterforskningsverktøyene gjennomgås nøye før bruk i litteraturen.[71]

En forutsetning for å gjennomføre grundige tester av programmene er at testerne har tilgang til programmets innhold, noe som ikke alltid er tilfellet.

2.1.4 Politiet har ikke tilgang til kodene, algoritmene eller treningsdataen modellen er bygget på

Politiet i dag bruker både «hyllevare» [72] og egenutviklede programmer. Egenutviklede programmer brukes oftest der én enkelt oppgave skal løses. For mer komplekse og fleksible løsninger benyttes hovedsakelig hyllevare. Leverandørene av slike produkter har liten interesse av å utlevere koder og treningsdata av flere grunner. For det første argumenteres det for at produktet er beskyttet av vern av forretningshemmeligheter.[73] For det andre vil utlevering av koder og treningsdata kunne avsløre svakheter i programmet som kan misbrukes. For det tredje vil avsløringer av potensielle svakheter gjøre at leverandører taper terreng i markedet. Det er derfor ofte i leverandørenes interesse å holde kortene tett til brystet. Mer om dette i neste punkt.

Programmene politiet bruker benytter har enten «åpen» kildekode eller «lukket» kildekode.[74] Åpen kildekode lar alle ha innsyn i programmets oppbygning, mens lukket kildekode skjuler programmets oppbygning. Begge har sine fordeler og ulemper.

Åpen kildekode lar politiet – og andre – føre kontroll med programmet og rapportere inn feil som kan fikses (såkalt «debugging»). Innsynet åpen kildekode tilbyr gjør at brukerne – herunder aktører i rettsvesenet – kan føre en bedre kontroll med programmets oppbygning, noe som igjen styrker bevisenes notoritet. På den andre siden vil det være lettere for aktører med uærlige hensikter å utnytte svakheter i systemer som åpenbarer seg gjennom kodene.

Lukket kildekode har den fordelen at programmene ikke er like utsatt for å eksponere svakheter i programmet som senere kan utnyttes. Det fører samtidig til at brukerne ikke vet hvordan programmet er bygd opp og muligheten for kontroll svekkes. Kodene er et

---

[71] Katrin Franke og André Årnes på side 336 i Årnes, A. mfl. *Digital Forensics*.
[72] Ikke egenutviklede programmer, men snarere produkter på markedet.
[73] Loomis vs Wisconsin avsnitt 51.
[74] Hentet fra Store Norske Leksikon 10. april 2020: «Åpen kildekode, viser til programvare som distribueres under forutsetning av at også kildekoden skal gjøres tilgjengelig for brukerne, i motsetning til programvare som utelukkende distribueres i binærform, og der kildekoden er opphavets strengt bevarte hemmelighet – i noen tilfeller også patentbeskyttet.»

markedsprodukt som i mange rettssystemer er beskyttet som forretningshemmeligheter. Når kodene til programmet holdes skjult gjør det at det blir vanskeligere å føre den nødvendige kontrollen med programvaren, som igjen kan gå på bekostning av programmets pålitelighet.

Et eksempel på et program som har lukket kildekode er Palantir, hvor politiet har kjøp leverandørens «Gotham»-løsning til 81 millioner kroner.[75] Katrin Franke, professor ved NTNU Gjøvik, stiller i et intervju til politiforum.no spørsmål om hvorfor Norge har valgt Palantir som leverandør når det finnes andre leverandører på markedet som opererer med «åpen kildekode».[76] Eksempelvis tilbyr programvaren «Sqrrl» tilsvarende løsninger, og har «åpen kildekode».

Franke stiller spørsmåltegn ved om man har sett for seg en situasjon der programmets pålitelighet blir tvistetema i retten og om påtalemyndigheten i den forbindelse kan føre tilstrekkelige beviser for at programmet fungerer nøyaktig slik det påstås. Det er lett følge Frankes tankegang her; hvordan kan man svare på kontrollerende spørsmål for et program man ikke har innsyn i?

Det er verd å merke seg at norsk politi kun har kjøpt «søkepakken» og hittil ikke har utvidet til noen av de mer sofistikerte analyseverktøyene Palantir tilbyr, eksempelvis «predictive policing». På den andre siden skal det finnes flere programmer i politiet som benytter lukket kildekode. Utfordringene rundt programmer bygget på lukket kildekode er i straffeprosessuell sammenheng lite problematisert. Fokus syns derimot først og fremst å ha vært rettet mot personvernspørsmål, se f.eks. Meld. St. 28 (2018-19) på side 61 der det uttales av Palantir-avtalen kan utfordre borgernes personvern.[77]

2.1.4.1 Leverandørenes interesser setter norsk politi i en ugunstig posisjon

Et stort problem er at leverandørene ikke vil selge de digitale løsningene sine med en «error rate» ettersom det kan skade deres konkurranseposisjon i forhold til andre selskaper. Det er lett å forstå at et program som selges med forbehold om feil selger dårlig. Det har derfor vært en tendens til å selge løsningene så lenge de fungerer. I likhet med brukeren – i dette tilfellet

---

[75] Norge signerte i 2009 Prüm-avtalen med EU og i 2012 signerte PCSCavtalen med USA. Avtalene sikrer Norge muligheten til å søke etter DNA og fingeravtrykk direkte i EUs og FBIs databaser.

[76] Artikkel på politiforum.no *«Palantir holder planene for framtiden hemmelige. Professor er bekymret for at data kan havne i private selskapers hender.»* Tilgjengelig på: https://www.politiforum.no/artikler/palantir-holder-planene-for-framtiden-hemmelige-professor-er-bekymret-for-at-data-kan-havne-i-private-selskapers-hender/434026 (sist lest 21.05.20)

[77] Meld. St. 28 (2018-2019) Datatilsynets og Personvernsnemdas årsrapporter for 2018. Side 61.

det norske politiet – vil det for leverandøren være tid- og ressurskrevende å kontrollere egne programmer med tanke på å avdekke feilrate. Det som favoriseres vil ofte være å utvikle programmet ytterligere. Resultatet ved at programmene selges uten et slik forbehold er at bevisene som disse programmene kan produsere får karakter av å være «autentiske», og man risikerer – dersom det de ikke kontrolleres – at de blir tatt for akkurat det. Med «autentisk bevis» menes bevis det ikke stilles spørsmål ved og som uten videre blir tatt for å representere faktiske forhold. Dette er helt klart uheldig. Som vist ovenfor må rettens aktører på en eller annen måte kunne vite noe om programmets overordnede pålitelighet om det skal benyttes på et så inngripende rettsområde som strafferetten, jf. beviskravet der «enhver rimelig tvil skal komme den tiltalte til gode».

Ansvaret for å vurdere påliteligheten av de digitale verktøyene skyves dermed i realiteten over på kjøperen, dvs. det norske politiet. Slik det straffeprosessuelle systemet er organisert i dag er det i hovedsak politiet som skal ha kompetansen om de metodene og de digitale verktøyene de benytter. Det er neppe urimelig å påstå at feil i de digitale verktøyene som politiet benytter vil være vanskeligere for en forsvarer eller dommer, å oppdage grunnet særkompetansen som trengs. Dersom eventuelle feil ikke oppdages i politiet risikerer man «følgefeil» gjennom resten av systemet.[78] Det må ut ifra disse betraktningene være ønskelig at politiet har tilstrekkelig kunnskap om verktøyene de benytter, slik at ansvaret for å oppdage eventuelle feil i for stor grad ikke skyves over på forsvarerne og dommerne.

3 Analyse av det straffeprosessuelle systemet og politiets praksis

I forrige kapittel ble de faktiske omstendighetene i tilknytning til KI-systemene trukket frem. Det ble også gjort noen korte betraktninger i tilknytning til forhold som vil ha betydning i en analyse av det straffeprosessuelle systemet. I dette kapittelet vil dagens rettstilstand bli analysert og evaluert.

Det tas utgangspunkt i de straffeprosessuelle grunnprinsippene og de underliggende hensynene de er ment å ivareta. Når det er gjort vil rettssikkerhetsmekanismene vurderes enkeltvis, og helhetlig. Det vil så oppsummeres og konkluderes om dagens bevissystem er adekvat til å møte KI-problematikken, de lege lata. I adekvatvurderingen vil vurderingstemaet

[78] Feilen forplanter seg på et tidlig stadium i rettsprosessen (etterforskningen) uten å bli oppdaget senere i prosessen.

være hvorvidt de grunnleggende hensynene i straffeprosessen ivaretas når bevis som er behandlet av kunstig intelligens føres for retten.

3.1 Straffeprosessens formål, hensyn og grunnprinsipper

Straffeprosessens formål er å fremme strafferettspleien.[79] Straffeprosessen skal sammen med strafferetten og straffegjennomføringsretten sikre at strafferettspleiens formål realiseres, dvs. at skyldige får sin straff samtidig som uskyldige går fri.[80]

Straffeprosessen kan beskrives som fremgangsmåten myndighetene må benytte når fysiske eller juridiske personer skal dømmes til straff. Straff er myndighetenes mest inngripende virkemiddel og er definert som *«et onde som staten tilføyer en lovovertreder på grunn av lovovertredelsen, i den hensikt at han skal føle det som et onde».*[81] Alvorligheten av straff som ondepåføring tilsier at prosessen som fører frem til domfellelsen må være gjort på en tillitsskapende og betryggende måte.[82]

Når myndighetene skal realisere denne målsetningen trenger den visse fullmakter for å gjøre inngrep i borgernes rettigheter. Da det kan være nødvendig å sette makt bak truslene om straff, er myndighetene tildelt et maktmonopol overfor borgerne.[83] Ved tildelingen av slike inngripende fullmakter taler gode grunner for at fullmaktene har legitimitet ved å være forankret i befolkningen.[84] For at det skal være konsensus i befolkningen om at fullmaktene er nødvendige er tilliten til rettssystemet essensielt. Det er ikke tilstrekkelig for rettsikkerheten at den materielle retten er ivaretatt dersom prosessen ikke er tillitsskapende.[85] Hensynet til tillit gjelder selvsagt for den tiltalte selv, men også for andre aktører i straffeprosessen og allmenheten for øring.[86]

---

[79] Øyen, Ø. 2. utg. 2016. *Straffeprosess.* Bergen: Fagbokforlaget. Side 24.
[80] Ibid.
[81] Johs Andenæs definisjon av straff får tilslutning av Høyesterett i Rt. 1977 s. 1207 (side 1209). Forståelsen er også tatt inn i forarbeidene til lov 20 mai 2005 nr. 28 om straff (straffeloven), Ot.prp. nr. 90 (2003-2004) på side 20.
[82] Øyen, Ø. *Straffeprosess*. Side 25.
[83] Med makt menes her fysisk makt med den hensikt å få gjennomført (realisert) den materielle retten.
[84] Forankringen skjer gjennom representativt demokrati.
[85] Rui, J. P. (2014*). Straffeprosessen i perspektiv*. Publisert i: Jussens venner. Juni 2014 (Volum 49). punkt 4.3.
[86] Ibid.

I vår rettstradisjon kreves det at systemet er effektivt, tillitsskapende, betryggende, ivaretar menneskeverdet og er kostnadseffektivt.[87] Hensynene vil ofte gå på bekostning av hverandre slik at de må avveies mot hverandre.

Konstitusjonelle rettigheter og menneskerettighetene er de mest fremtredende rettighetene på straffeprosessens område. Grunnlovens rang og lov 21 mai 1999 nr. 30 om styrking av menneskerettighetenes stilling i norsk rett (Menneskerettloven) § 2, jf. § 3. gir reglene en metodisk særstilling i norsk rett. De viktigste kan sies å være Grl. § 96, «ingen skal straffes uten dom.» Det samme framkommer av EMK art. 6, 1. første punktum. Hensynet bak § 96 er hovedsakelig at uskyldige ikke skal straffes. I Grl. § 95, EMK art. 6 og SP art. 14 framkommer det at alle skal ha en rettferdig rettergang.[88] Etter Grl. § 96 annet ledd og EMK art. 6 nr. 2 og 3 skal alle anses uskyldig inntil det motsatte er bevist etter loven. I denne uskyldspresumsjonen kan det strenge beviskravet forankres.[89] *«Enhver rimelig tvil om de faktiske forholdene skal komme den siktede til gode»*, jf. Rt. 2005 s. 1353 avsnitt 11.[90]

Et annet grunnleggende hensyn er hensynet til rettssikkerhet. Utrykket er gjennom årene blitt utvannet og har ikke et konkret innhold.[91] Etter Jon Petter Ruis forståelse refererer begrepet i straffeprosessen først og fremst til kravet om et «materielt riktig resultat».[92] For å oppnå et materielt riktig resultat må straffeprosessens regler sikre at den «materielle sannhet»[93] blir kartlagt. Den «materielle sannhet» søkes i norsk prosessrett gjennom bevisføring for ulike bevistemaer og det eksiterer et alminnelig prinsipp om forsvarlig og rettferdig saksbehandling, jf. Rt. 2006 s. 856 avsnitt 30. En godt opplyst sak gir de beste forutsetningene for å avgjøre spørsmålet om straff, og motsetningsvis, om den straffeforfulgte skal frifinnes. Regler som er ment å ivareta disse interesse kalles rettssikkerhetsmekanismer og skal bidra til at uskyldige ikke dømmes. Reglene eller «mekanismene» anses som

---

[87] Øyen, Ø. *Straffeprosess*. Side 26.
Det er på det rene at hensynene i straffeprosessen aldri kan etterleves 100 % ettersom de begrenses av blant annet prosessøkonomiske hensyn, menneskelige faktorer som gjør full objektivitet umulig, etc.
[88] UNs International Covenant on Civil and Political rights, 16.12.1966. (Den internasjonale konvensjonen om sivile og politiske rettigheter, heretter SP).
[89] NOU 2016: 24 Ny straffeprosesslov. Kapittel 13.2 om alminnelige bevisregler. Gjennomgang av gjeldende rett. I første setning på side 263 uttales det: «I lang tid hvilte det strafferettslige beviskravet alene på sedvanerett, men etter grunnlovsendringene i mai 2014 har den også forankring i uskyldspresumsjonen i Grunnloven § 96 annet ledd.»
[90] Tilsvarende Rt. 2009 s. 1109 avsnitt 39-40 og HR-2011-1969-U avsnitt 46.
[91] Jon Petter Rui, *straffeprosess i perspektiv.* Punkt 4.2.
[92] Ibid.
[93] Ruis presisering av «sannhet», ibid: Med «sannhet» kan ikke forstås den absolutte objektive sannhet i ytterste empiristiske, erkjennelsesteoretiske konsekvens, à la Descartes' «jeg tenker altså er jeg». Et slikt nivå av sannhet er stort sett umulig å nå i de fleste praktiske sammenhenger.»

essensielle i en rettstat. Å straffedømme uskyldige anses som en av de verste feilene et samfunn kan gjøre overfor et individ.[94] I den forbindelse sies det at det er bedre å la 10 skyldige gå fri, enn at én uskyldig dømmes.

Samtidig som staten skal verne borgernes rettigheter i straffeprosessen, er samfunnet helt avhengig av at myndighetene til enhver tid kan drive effektiv kriminalitetsbekjempelse. Hver enkelt persons rettigheter kan ikke trekkes så langt at straffelovgivningen ikke blir realisert. Et dysfunksjonelt rettsapparat vil gå ut over tilliten til statens strafferettspleie, herunder straffeprosessen. Hensynet til et effektivt rettsvesen står derfor sterkt og vil på mange måter være en forutsetning for straffeprosessen.

Objektivitetsplikten i politiet og påtalemyndigheten skal sikre en objektiv etterforskning etter strpl. § 55 a fjerde ledd og § 226, mens domstolens uavhengighet skal sikre en rettferdig bedømmelse av bevisene, jf. Grl. § 95. Disse grunnleggende hensynene bidrar til å fremme et materielt riktig resultat og tillit.

Hovedregelen for bevisføring i norsk straffeprosess er det ulovfestede prinsippet om fri bevisføring.[95] Prinsippet innebærer at partene i utgangspunktet kan føre de beviser de ønsker.[96] Hensynet som begrunner regelen er en antagelse om at partene vil bidra til sakens opplysning dersom de får anledning til å føre de beviser de ønsker, samt at en bedre opplyst sak i utgangspunktet gir et bedre avgjørelsesgrunnlag enn en mindre opplyst sak.[97]

I selve bevisvurderingen gjelder det ulovfestede prinsippet om fri bevisvurdering.[98] Prinsippet innebærer at det i utgangspunktet ikke legges noen føringer på bevisbedømmeren for hvordan vurderingsmetode som skal benyttes eller hvordan bevisverdien skal fastsettes.[99] Prinsippet baserer seg på en antagelse om at sannheten gjennomgående fremmes i størst grad hvis

---

[94] Op.cit, punkt. 4.3
[95] Rt. 1990 s. 1008 på s. 1110. Se også Rt. 2008 s. 605 avsnitt 13, Rt. 2005 s. 1353, avsnitt 13, Rt. 2002 s. 1744 på s. 1746, Rt. 1996 s. 1114 på s. 1116 og Rt. 1994 s. 610 på s. 614.
[96] Unntak i strpl. §§ 134, 292, 293, 295 og 301.
[97] NOU 2016: 24 Ny straffeprosesslov. Side 257. Tredje avsnitt i punkt 13.2.3.
[98] Prinsippet er ble tatt ut av gjeldende straffelov da prinsippet ble ansett som selvsagt. Se NUT 1969: 3 Innstilling om rettergangsmåten i straffesaker fra Straffeprosesslovkomiteen (Komiteen til revisjon av straffeprosessloven). Side 308 under «Til § 311. [lovens § 305]». Prinsippet kan imidlertid gjøre inntog i den nye straffeprosessloven, se NOU 2016: 24 side 573 under punktet «Til § 7-5. Bevisvurdering og beviskrav».
[99] Anders Løvlie på side 538 i Hedlund, M. mfl. *Bevis i straffesaker: utvalgte emner*.

bevisbedømmelsen er fri.[100] Bedømmelsen er imidlertid ikke helt fri, det stilles krav til rasjonalitet og samvittighetsfullhet.[101]

I avhandlingen fokuseres det på de bevis som føres for retten i spørsmålet om straff og som er produsert av et dataverktøy som betegnes som kunstig intelligens. Typene bevis kommer i to former. Tradisjonelle fysiske bevis som har blitt behandlet av kunstig intelligens (analyser av fingeravtrykk, lyder, stemmer, video, etc) og digitale bevis som også har vært behandlet av kunstig intelligens (digitale spor som lokasjonsdata, IP-adresser, etc).

I den forbindelse er det programvarens «out-put» som det trekkes slutninger fra. Sagt med andre ord, programmets «out-put» fungerer som «bevisdata» i bevisbedømmelsen.[102] Allment anerkjente erfaringssetninger er supplementet som gir beviset mening.[103] En variabel som alltid har betydning for bevisvurderingen er hvordan bevisbildet for øvrig ser ut i saken. Det er den objektive gjerningsbeskrivelsen som helhet som må bevises ut over enhver rimelig tvil og ikke hvert enkelt bevis, eller forhold som ikke har betydning for gjerningsbeskrivelsen.[104] Bevis produsert av kunstig intelligens vil i de fleste tilfeller inngå i en beviskjede snarere enn å fungere som det eneste fellende beviset i saken. I så tilfelle kan potensielle feilkilder ved KI-beviset neglisjeres på en helt annen måte (gitt at det øvrige bevisbildet er informativt og robust), ettersom bevisbedømmeren vurderer bevisbildet som helhet. For at oppgaven skal ha mest for seg problematiseres imidlertid de tilfellene der beviset har en så sentralt funksjon at spørsmålet om pålitelighet blir avgjørende for spørsmålet om straff.

3.2 Utgangspunkt for analysen

Rettssikkerhetsmekanismene som skal sikre at feilaktige bevis ikke legges til grunn har tradisjonelt vært: (1) politiet og påtalemyndighetens ansvar for et forsvarlig avgjørelsesgrunnlag, (2) retten til innsyn og kontradiksjon, samt (3) rettens ansvar for et forsvarlig avgjørelsesgrunnlag. Det er disse rettighetene og pliktene som skal drøftes.

---

[100] NOU 2016: 24 side. 262.
[101] Andenæs, J. og Myhrer, T. 4. utg. 2009. *Norsk straffeprosess*. Oslo: Universitetsforlaget. Side 166.
[102] Løvlie, A. *Rettslige faktabegreper.* Side 162. Løvlie beskriver «bevisdata» som «det bevisene i seg selv uttrykker». Bevisdatabegrepet har dermed et bredt nedslagsfelt og kan beskrives som «rådata».
[103] Løvlie, A. *Rettslige faktabegreper.* Side 155, punkt 4.3.2.2.
[104] Det må også føres tilstrekkelige bevis for subjektiv skyld.

Det sees følgelig bort fra reglene om bevisavskjæring. Jeg kan ikke se at et bevis som er produsert av kunstig intelligens kan avskjæres på det grunnlag alene.[105] Det sees og bort fra bevis som er innhentet på ulovlig eller utilbørlig vis.

En potensiell løsning av problematikken som er benyttet i andre rettssystemer er utarbeidelsen av bevisvurderingsregler fra domstolene, eksempelvis USA. Dette var tilfellet i Loomis vs Wisconsin. Retten i Wisconsin oppstilte strenge opplysningskriterier som måtte følge med beviset dersom det skulle føres.[106] Eksempelvis måtte retten på forhånd opplyses om at leverandørene ikke ønsket å dele informasjon om hvordan KI-systemet fungerte, hvilket datasett programmet hadde trent på, at tidligere studier hadde konkludert med skjevheter i programmet, samt et krav om at systemet kontinuerlig måtte re-evalueres. Loomis anket, men Wisconsin Supreme Court avviste saken. Løsningen i møte med problematikken ble altså å benytte domstolen prejudikatskraft til å legge føringer på hvordan bevisene skulle legges fram og hvordan det skulle bedømmes. Fordelen med denne rettssikkerhetsmekanismen er at den er ment å ivareta rettssikkerheten på et høyere nivå enn fra sak til sak. Svakheten til dommen er at prejudikatet bare gjelder COMPAS og nært beslektede programmer.

Det er imidlertid lite sannsynlig at problemet kan løses på samme måte i Norge grunnet Høyesteretts begrensede kompetanse til å legge føringer på bevisbedømmelsen. Tanken om at det ikke kreves noen særlige kvalifikasjoner for å bedømme bevis har stått stødig i vår rettstradisjon, jf. lekdommer og juryordningen. Fagdommerne i Høyesterett har sterk juridisk bakgrunn, men det impliserer ikke at de har særlig egenskaper for bevisbedømming.[107] Det kan vanskelig tenkes at norsk Høyesterett kan legge strenge prejudikatsregler på føringen og bedømmingen av bevis, slik som domstolen i Wisconsin. Merk at i Wisconsin-saken ble KI-verktøyet benyttet i spørsmålet om straffeutmåling og ikke i spørsmålet om straff. Det må antas at det ville blitt stilt minst like strenge krav dersom et KI-verktøy med base i maskinlæring hadde presentert et bevis i spørsmålet om straff.[108] I de sakene lagmannsretten har tydd til generelle prinsipper i bevisbedømmelsen har Høyesterett opphevet etter prinsippet

[105] Bevisavskjæringsreglene gir en snever adgang til å avskjære bevis. Eksempelvis kan ikke et bevis avskjæres fordi det antas å ha lav beviskraft.
[106] Loomis vs Wisconsin. Se side 763–65.
[107] Løvlie, A. *Rettslige faktabegreper.* Side 295.
[108] Det anses som mer prekært for rettssikkerheten å stå overfor spørsmålet om straff enn om straffeutmålingen.

om fri bevisbedømmelse, se f.eks. Rt. 2001 s. 543 (side 545).[109] Tanken om at KI-problematikken kan løses gjennom konkrete prejudikatssetninger er derfor fjerntliggende.

En mer indirekte tilnærming fra Høyesterett kan også nevnes. Høyesterett har gjennom «teknisk pregede erfaringssetninger» prøvd underrettenes bevisbedømmelse og opphevet avgjørelser på grunnlag av *«generelle erfaringssetninger eller andre generelle vurderingsgrunnlag»*, jf. Rt. 1984 s. 840 (side 842).[110] Særlig har dette vært aktuelt for erfaringssetninger knyttet til kjøring i alkoholpåvirket tilstand. Virkemidlet består i å oppheve lagmannsrettens dom ved å fastslå at begrunnelsen ikke i tilstrekkelig grad bygger på teknisk pregede erfaringssetninger.[111] For det tilfelle at et KI-bevis ikke utredes tilstrekkelig på bakgrunn av de tekniske erfaringssetningene vil muligens Høyesterett oppheve en eventuell dom fra lagmannsretten på dette grunnlag, og på den måten forme hva som må belyses. På den andre siden er virkemiddelet bare en måte å stille krav til forholdet mellom premisser, informasjonsgrunnlag, og konklusjon.[112] Selve informasjonsgrunnlaget prøves ikke, jf. Rt. 2012 s. 897 avsnitt 14.[113] Siden det langt på vei er informasjonsgrunnlaget som utgjør utfordringene i KI-problematikken vil heller ikke dette virkemidlet fra Høyesterett være egnet til å få bukt med problematikken. Unntaket kan selvsagt være den enkelte sak som ankes til Høyesterett på det grunnlag som er nevnt, men prejudikatsvirkningen av opphevelsen vil ikke være kraftig – eller konkret – nok til å legge nødvendige føringer på bevisførselen og bedømmelsen. Det er derfor ikke en egnet løsning å overlate til Høyesterett å løse utfordringene knyttet til KI-systemer gjennom sin mulighet til å oppheve lagmannsrettens dommer på bakgrunn av utilstrekkelig skriftlig begrunnelse.

Domstolens funksjon vil i møte med problematikken begrense seg til den enkelte sak gjennom strpl. § 294, plikten til å sørge for sakens opplysning. Plikten behandles nedenfor.

3.2.1 Kunstig intelligens stilling i bevisspørsmål i straffeprosessen

Avhandlingen skal analysere problemer på to plan: Det første er der det framkommer at beviset er produsert av et KI-system, f.eks. av politirapportene. Det andre er der det ikke framkommer at beviset er produsert ved bruk av kunstig intelligens. Grunnen til det skal

---

[109] Nederst på side 545: «Om det er gjort sannsynleg, må prøvast etter prinsippet om fri bevisvurdering. Det kan ikkje tolkast inn i § 459 særlege prinsipp for vektlegging av bevis.»
[110] Løvlie, A. *Rettslige faktabegreper.* Side 503.
[111] F.eks. Rt. 1985 s. 168, Rt. 1983 s. 1275 og Rt. 1982 s. 1264.
[112] Løvlie, A. *Rettslige faktabegreper.* Side 505.
[113] Avsnitt 14: «Når lagmannsretten fant å kunne legge blodprøveresultatet til grunn, er det ingen mangel ved domsgrunnene at den ikke tok stilling til hvor mye alkohol A hadde drukket.»

knyttes noen bemerkninger til sistnevnte tilfelle er at dagens politipraksis ikke rutinemessig opplyser om at kunstig intelligens har vært involvert. Det ligger således an til at problematikken aldri adresseres i rettssystemet, noe som er uheldig, jf. de potensielle feilkildene slike systemer operer med. At det ikke rutinemessig opplyses om bruken kan skyldes flere forhold. Politiet har for eksempel kjøpt et digitalt hjelpemiddel uten å vite hvordan det helt konkret fungerer samtidig som de ikke har innsyn i programmets oppbygning. Informasjonsflyt kan være et annet problem innad i politiet. Det kan også være tilfelle at de som kjøper programvaren har adekvat informasjon om programmets styrker og svakheter, herunder at programmet benytter avanserte maskinlæringsteknikker, men at det aldri når etterforskerne.

Dersom informasjonen ikke når rettens aktører[114] vil de være uvitende om de særskilte utfordringene som kan følge bevis som er produsert av kunstig intelligens, sammenlignet med mer tradisjonelle digitale systemer. Med rettens aktører menes i denne oppgaven dommerne, forsvarssiden og aktor. Dersom rettens aktører er uvitende om bruken av KI-systemer og omstendighetene rundt, vil heller ikke rettssikkerhetsmekanismene komme i spill, noe som i utgangspunktet kan være et stort rettssikkerhetsmessig problem, jf. hensynet til det materielle sannhets prinsipp. Utfordringene kan i ytterste konsekvens komme i direkte konflikt med kravet til forsvarlig avgjørelsesgrunnlag, og hensynet til et materielt riktig resultat. Konsekvensene kan gå i begge retninger gjennom falske positive og falske negative resultater. Resultatet kan bli at skyldige går fri og at uskyldige blir dømt, hvorav sistnevnte er mest alvorlig. Et spørsmål som må belyses er hvorfor denne informasjonen ikke når rettens aktører.

Hvorvidt hver enkelt dataetterforsker eller etterforsker vet at programmet de bruker er et KI-system har i utgangspunktet begrenset betydning, så lenge informasjonen ikke når påtaleansvarlig eller aktor.

Dagens politipraksis er at etterforskere utarbeider en dataetterforskningsrapport som sendes til påtaleansvarlig da det er påtaleansvarlig som avgjør spørsmålet om tiltale. Riksadvokaten uttalte i 2019 at *«[k]ravene til dokumentasjon og notoritet er i dag meget strengere enn bare for få år siden.»*[115] Dataetterforskningsrapporten utarbeides for å sikre notoriteten til beviset, dvs. at etterforskningskriteriene skal være etterprøvbare. Rapporten har som formål i sikre at

[114] I oppgaven forstås rettens aktører som dommerne, forsvarssiden og aktor.
[115] Riksadvokatens notat (07.01.19) *Notat om utviklingen ved etterforskningsfelt.* På side 5, punkt 3.4.4 om dokumentasjonskrav.

bevisets integritet er ivaretatt og at dette er dokumentert.[116] Rapporten har på den måten funksjon av å være hovedredskapet for å dokumentetere hvordan politiet har gått frem, og om beviset er forsvarlig behandlet. Dataetterforskningsrapportene har imidlertid møtt kritikk av flere grunner, blant annet fordi de ikke er nøyaktige nok og dermed åpner for antagelser og feiltolkninger. [117]

Dagens praksis i politiet skaper utfordringer når det rutinemessig bare føres opp navnet på dataverktøyet som har blitt benyttet, mens informasjonen om at beviset er behandlet av et KI-system utelates. Som vi har sett ovenfor er det generelt sett en ikke ubetydelig forskjell på usikkerhetsmomentene for KI-programmer sammenlignet med mer tradisjonelle dataprogrammer. De mer tradisjonelle dataetterforskningsprogrammene har gjerne blitt brukt til oppgaver som speilkopiering av digitale enheter, uthenting tekstmeldinger, logger av nettsurfing, etc., en langt mindre kompleks oppgave enn det dagens verktøy er kapable til.[118] Fra et rettssikkerhetsperspektiv vil det ha formodningen for seg å opplyse om at kunstig intelligens har behandlet beviset av hensyn til kontradiksjon, og gjennom det, sakens opplysning. Hensynet til sakens opplysning baserer seg på forutsetningen om at all informasjon som kan ha betydning for saken skal legges fram. Hensynet til bevisets integritet taler derfor klart for at det burde informeres om at kunstig intelligens har behandlet bevisene mot den tiltalte, jf. at unøyaktighetene som kan følge KI-systemer i utgangspunktet er for stor til å neglisjeres.[119]

I dagens politipraksis ligger det altså til rette for at bevis som er produsert av KI-systemer ikke når aktørene i rettssystemet, noe som i utgangspunktet er uholdbart dersom man ønsker å holde samme høye standard som for øvrige digitale bevis[120] Politiet har en plikt til å behandle bevis på en forsvarlig måte og ansvaret reguleres ofte av strenge retningslinjer for å bevare bevisets integritet.[121] Systembetraktninger av de ulike formene for forsvarlig behandling av

---

[116] André Årsnes på side 54 i Årnes, A. mfl. *Digital Forensics*.
[117] Erlandsen, T-E. (2019). *Fallacies when Evaluating Digital Evidence Among Prosecutors in the Norwegian Police Service.* (Masteroppgave, NTNU). Tilgjengelig på: https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2617771
[118] Tradisjonelle dataetterforskningsprogrammer følger gjerne er mer tradisjonell «steg for steg» modell, sammenlignet med avanserte maskinlæringssystemer. De mer tradisjonelle verktøyene kan for eksempel hente ut tidsstempler for meldinger eller nettlogger. Som vist i op.cit (Erlandsen), kan slike programmer utgjøre feilkilder, enten via uthenting eller via tolkning av dataene.
[119] Kapittel 2.
[120] Riksadvokatens rundskriv nr. 3/2018. (Revidert 21.02.19) *Kvalitetskrav til straffesaksbehandlingen i politiet og ved statsadvokatembete mv. (kvalitetsrundskrivet).* Side 28. Punkt 4.11.2.
[121] Inger Marie Sunde på side 600 i Hedlund, M.A. mfl. *Bevis i straffesaker: utvalgte emner.*

bevis og et generelt ønske om usikkerhetsminimalisering, taler for at rettens aktører meddeles at kunstig intelligens har vært involvert i bevisbehandlingen.

I samme retning trekker hensynet til tiltaltes rett til å kunne imøtegå beviset.[122] Å føre et bevis den tiltalte vet er uriktig eller unøyaktig, vil rokke ved den straffeforfulgtes tillit til rettssystemet. I et slikt scenario er det lett å ha sympati med den tiltalte som åpenbart vil føle seg urettferdig behandlet. Uten ytterligere informasjon vil det imidlertid være vanskelig for den tiltalte å sannsynliggjøre at det hefter feil ved beviset.[123]

Utover dataetterforskningsrapporten finnes det ikke retningslinjer eller rutiner som er ment å viderebringe slik kunnskap til rettens aktører. Resultatet er at det ikke informeres rutinemessig om at KI-systemer har vært involvert i behandlingen av beviset. Dersom informasjonen mot formodning skulle nå rettens aktører vil det skyldes enkeltindividers initiativ, eksempelvis en dataetterforsker som sitter på kunnskapen og velger å opplyse om det i rapporten. Det kan samtidig tenkes at en usedvanlig teknisk advokat mistenker at kunstig intelligens har vært involvert og kommer med innsigelser. Det kan imidlertid ikke sies å være en langsiktig løsning å overlate problematikken til enkeltindivider.

Problemområdet i politiets og påtalemyndighetenes rutiner er dermed identifisert. Dataetterforskningsrapporten som skal ivareta hensynet til notoritet, bevisets integritet, en tillitsvekkende prosess, skriftlighet og kontradiksjon, inneholder ikke den nødvendige informasjonen som kan være avgjørende for å imøtegå beviset.

Siden det er politiet som benytter disse etterforskningsmetodene er det også i politiet kompetansen er ment å ligge. Det vil da neppe være urimelig å påstå at dersom ikke politiet oppdager en eventuell feil vil oppgaven bli desto mer utfordrende for rettens øvrige aktører. Det er derfor særlig viktig at problemet adresseres der ekspertisen er ment å ligge.

Det tas derfor til ordet for at politiets praksis på dette området må endres.

Når utfordringene rundt bevisbehandlingssystemer med base i kunstig intelligens blir kjent vil det utløse en plikt for politiet og påtalemyndigheten til å rette opp i ubalansen, dersom beviset ennå ønskes ført. Dersom bevisene – etter at utfordringene er allment kjent – fortsatt føres

---

[122] Retten til kontradiksjon er et grunnleggende straffeprosessuelt prinsipp og er inntatt i Grl. § 95 om retten til en rettferdig rettergang og EMK art. 6 nr. 3.
[123] Andre beviser den tiltale kan sitte på vil selvsagt kunne motbevise KI-systemets «out-put», eksempelvis vitneobservasjoner på at den tiltalte var et annet sted enn det KI-systemets «out-put» antyder.

uten tilstrekkelig informasjon eller forbehold, står påtalemyndigheten i fare for å utfordre prinsippet om «equality of arms».[124]

For dagens situasjon må konklusjonen bli at grunnleggende straffeprosessuelle hensyn ikke er ivaretatt ettersom potensielt viktig informasjon om KI-bevis ikke når rettens aktører. Hensynet til det materielle sannhets prinsipp utfordres jevnlig ved at kravet til forsvarlig behandling av bevis, herunder sikre bevisets integritet og saksopplysning, ikke overholdes.

Prosessen er på bakgrunn av de faktiske forholdene ikke tillitsskapende ettersom informasjonen om at kunstig intelligens har vært involvert, potensielt kan være avgjørende for på å påvise feil ved beviset. Effektivitetshensyn taler smått imot en utvidet kontroll, men kan ikke tillegges avgjørende vekt. Det er viktig å påpeke at objektivitetsplikten ikke anses brutt ettersom dette er en nokså ukjent problemstilling. Selv om politiet og påtalemyndigheten subjektivt går inn for å sikre et forsvarlig avgjørelsesgrunnlag er ikke dagens system tilrettelagt for å reelt sett overholde kravet. I den forbindelse vil det bemerkes at oppgaven ikke har til hensikt å beskylde politiet for å aktivt holde tilbake potensielt viktig informasjon. Det må antas at når utfordringene blir allment kjent – andelen av KI-verktøy øker, og problematikken gjennomgås, vil praksis endres. Det konkluderes likevel med at dagens situasjon er uholdbart ut ifra et rettssikkerhetsperspektiv.

Analysen går nå over til oppgavens andre plan: for det tilfelle at KI-bruken blir kjent for rettens aktører. Under slike omstendigheter vil de rettssikkerhetsmekanismer som er ment å sørge for et forsvarlig avgjørelsesgrunnlag komme i spill. Det skal drøftes hvordan disse står seg mot utfordringene som kan følge bevis produsert av et KI-system.

3.2.2 Innsynsreglene

En eventuell mulighet for å føre kontroll med KI-problematikken kan være den straffeforfulgtes innsynsrett. Den straffeforfulgte har rett til innsyn i «sakens dokumenter» under etterforskningen etter stprl. § 242 og under forberedelsen av hovedforhandlingen, jf. strpl. § 264. Med en utstrakt innsynsrett vil den tiltalte og hans forsvarer rent faktisk kunne gjennomgå koder, vurdere treningsdata og teste programmet for selv å finne eventuelle

[124] Prinsippet ««equality of arms» innebærer at påtalemyndigheten og tiltalte skal gis de samme våpnene i møte med retten, herunder rett til all informasjon som kan ha betydning for saken. Se f.eks. Jasper v. United Kingdom, Grand Chamber 16 February 2000, (Saksnummer 27052/95). Avsnitt 51.

svakheter ved beviset. Imidlertid kan ikke dette anses som en særlig god løsning av følgende grunner:

(1) For det første er det usikkert om koder og treningsdata kan innfortolkes i «sakens dokumenter» rent juridisk. Det som tradisjonelt har inngått i «sakens dokumenter» er dokumenter som har blitt til eller framkommet under etterforskningen, jf. Rt. 2010 s. 655 avsnitt 40. En forutsetning for innsyn er også at politiet og påtalemyndigheten selv har tilgang til dokumentene det ønskes innsyn i. Som vist i kapittel 2 er ikke dette alltid tilfelle.

(2) For det andre vil innsyn i koder og treningsdata være beskyttet av vernet mot forretningshemmeligheter. Programvaren er et produkt og vil langt på vei nyte beskyttelse av vernet. Eksponering av programmets oppbygning kan avsløre svakheter som senere kan utnyttes.

(3) For det tredje vil innsyn i treningsdata kunne reise personvernspørsmål, eksempelvis dersom treningsdataen er biologiske eller biometriske data.

(4) For det fjerde vil effektivitetshensyn tale sterkt imot en slik løsning. En rett til å få utlevert denne informasjonen fører også med seg en rett til å ha adekvat tid til å forberede et forsvar, jf. EMK art. 6. nr. 3 bokstav b. Rettssystemet har behov for å få avgjort saker innen rimelig tid for å fungere på en større skala. Dersom man skulle basere seg på en løsning der forsvarssiden måtte gjennomgå KI-materien for å finne svakheter, ville tilliten til rettssystemet svekkes betraktelig, både på grunn av den uforholdsmessige tidsbruken før sakens ble avgjort, og belastningen det vil føre med seg for de involverte. I tillegg skyves ansvaret for kontrollen – under en slik ordning – i realiteten over på den tiltalte. Løsningen harmoniserer dårlig med forståelsen om at det er staten som besitter det overveldende antallet ressurser. På den andre siden er det ikke ønskelig å avskjære KI-beviset eller la det manglende innsynet gå på bekostning av beviskraften til KI-bevis, da myndighetene har behov for å sikre en effektiv kriminalitetsbekjempelse opp mot et dynamisk kriminalitetsbilde. I den forbindelse er bevis som er behandlet av kunstig intelligens en viktig brikke å kunne bekjempe moderne kriminalitet. For store inngrep i politiets mulighet til å benytte bevis som er behandlet av KI-systemer vil kunne føre til et dysfunksjonelt rettsapparat, særlig i de sakene der det kun er bruken av KI-systemer som kan bidra til at saken oppklares. Slike tilfeller kan blant annet være der mengden informasjon er overveldende, for eksempel ved et datainnbrudd/cyberangrep. Politiets tillit i samfunnet hviler i stor grad på dens evne til å bekjempe kriminalitet effektivt.

Å gi den tiltalte innsyn i koder og treningsdata møter flere faktiske og rettslige utfordringer, og vil gå på bekostning av grunnleggende straffeprosessuelle hensyn som effektivitet og tillit. Det er derfor klart at innsynsreglene ikke utgjør en adekvat rettsikkerhetsmekanisme i møte med KI-problematikken som er beskrevet ovenfor i kapittel 2.

3.2.3 Kontradiksjon

For det tilfellet at forsvareren får opplyst at programmet som har produsert beviset mot hans klient er et KI-system (muligens en KI med base i maskinlæring), vil forsvareren være bedre stilt til å ivareta klientens interesser enn hvis han blir uvitende.[125] Dersom forsvarssiden er uenig i beviset som er fremskaffet av et KI-system (eller bare ønsker å svekke det) kan forsvareren fremme innsigelser mot beviset, herunder at det er en «falsk positiv». Kontradiksjonsretten er et grunnprinsipp i straffeprosessen og kan forankres i EMK art. 6 og er inntatt i Rt. 2003 s. 1682 og Rt. 2007 s. 1255, begge i avsnitt 15.

Vestby og Vestby har i sin artikkel «Machine Learning and the Police: Asking the Right Questions» vist at det er mulig for «ikke-eksperter» å stille relevante og kritiske spørsmål om maskinlæringsalgoritmer og deres begrensinger.[126] For eksempel kan en forsvarer spørre om: Hvilken «in-put» data er brukt? Hvilken data er brukt til å trene algoritmen? Hvilken data er brukt som testdata? Hvor og når ble dataen samlet inn? Finnes det kjente variabler og i så fall hvordan veies de mot hverandre? Er treningsdataen samlet inn med det formål å brukes i modellen eller det er fragmenter av data som ble brukt til andre formål?

Siden påtalemyndigheten har bevisbyrden må de imøtegå eventuelle innsigelser fra forsvareren og svare tilfredsstillende på spørsmålene for at bevisverdien ikke skal svekkes. Dette har tradisjonelt ført til at det argumenteres over bevis og deres pålitelighet. Argumentasjonsprosessen anses langt på vei sannhetfremmende da den er med på å skape klarhet i relevante forhold.[127] Tradisjonelt har kontradiksjon rundt bevisene vært en god måte å skape en forståelse for bevisets pålitelighet.[128] Noe mer utfordrende har det imidlertid blitt etterhvert som bevisene er blitt mer og mer tekniske. KI-problematikken bringer samtidig med seg noen særskilte problemstillinger som rettssystemet ennå ikke har adressert. Det mest

---

[125] Som vist innledningsvis legger dagens politipraksis opp til at aktørene forblir uvitende, og konsekvensen vil antagelig være at ingen av rettssikkerhetsmekanismene kommer i spill for å adressere problemet. Dette er i og for seg uheldig med tanke på hva som står på spill for den tiltalte.
[126] Vestby, A. og Vestby, (2019). J. *Machine Learning and the Police: Asking the Right Questions*. Publisert i: Policing: A Journal of Policy and Practice. Juni 2019.
[127] En øvre grense finnes der bevisbildet og argumentene blir overveldende. Under slike omstendigheter vil ikke mer informasjon bidra til å fremme sannheten, men snarere virke forvirrende.
[128] Løvlie, A. *Rettslige faktabegreper.* Side 310.

åpenbare er «svart boks»-problematikken der hvert enkelt resultat ikke lar seg forklare. Spørsmålene må dermed rettes mot beviset på en indirekte måte, nemlig gjennom KI-systemets generelle pålitelighet, jf. kapittel 2.

Forsvarerens oppgave vil gjennom kontradiksjonen rundt beviset være å stille spørsmål som kan avsløre svakheter rundt beviset slik av bevisbedømmeren hensyntar svakhetene i sin bedømmelse. Forsvarerens oppgave vil – med hensyn til strafferettens beviskrav – ikke være å fjerne all tvil rundt beviset, men nok til at det hefter «rimelig tvil» ved det, gitt at det er det «fellende beviset» i saken.

Den faktiske utfordringen som setter begrensninger for kontradiksjonens rolle som sannhetsfremmende mekanisme er at politiet ofte ikke har tilgang til den informasjonen forsvareren spør etter. Herunder hvordan programmet er trent opp, hvilke mål programvaren har blitt instruert til å optimalisere, etc. Årsaken er som nevnt at politiet benytter seg av produkter som er innkjøpte og baserer seg på lukket kildekode. Man havner da i en situasjon der et av politiets digitale KI-systemer har gitt et «out-put» som ikke lar seg forklare, samtidig som retten er klar over de generelle utfordringene rundt KI-systemer. Konsekvensen er at bevisets bevisverdi svekkes. Dette er i og for seg positivt for den tiltalte.

På den andre siden er det nødvendigvis ikke en god løsning å plassere borgernes rettssikkerhet på skuldrene til den enkelte advokat, noe som kan være uheldig av flere grunner. For det første vil antagelig ikke KI-beviset bli tvistetema i alle de tilfellene det burde vært det. Det kan påstås at det straffeprosessuelle systemet dermed legger risikoen for oppdagelse over på forsvarssiden, noe som i utgangspunktet er uheldig da det går ut over tilliten til rettssystemet. Høyesterett har blant annet i Rt. 2007 s. 10 i avsnitt 16 trukket fram at når det opereres med usikkerhet rundt digitale verktøy vil hensynet til notoriet, rettssikkerhet og tillit stå sentralt. Hensynene Høyesterett trekker fram taler for at politiet selv må bære risikoen for usikkerheten så langt det er rimelig. For det andre vil straffeprosessens effektivitet lide under en omfattende gjennomgang av alle bevis som er behandlet av et KI-system.

For det tilfelle at kontradiksjonen avdekker at politiet ikke vet hvordan programmet er bygd opp vil det være interessant å undersøke de kryssende hensynene. På den ene siden vil bevisets pålitelighet ikke være like sterkt, på den andre siden er interessen påtalemyndigheten har av å føre det. Referansepunktet i drøftelsen vil være det strenge beviskravet i strafferetten.

På den ene siden trenger politiet digitale verktøy for å etterforske lovbrudd, samtidig som det ikke kan forventes at de utvikler alle de digitale hjelpemidlene selv. Det kan også argumenteres for at dersom bevisverdien til beviset vesentlig svekkes – grunnet utilstrekkelige svar på spørsmålene rundt KI-systemet – kan det føre til uriktige frifinnelser, noe som på sikt vil svekke tilliten til rettsvesenet. På den andre siden står borgernes rett til en rettferdig rettergang som omfatter retten til at rimelig tvil om faktiske forhold skal komme den tiltalte til gode, jf. at bevisgrunnlaget må være tilstrekkelig robust.[129] Samtidig kan det tenkes at bruken av programvaren er den eneste måten å etterforske akkurat denne typen kriminalitet på, for eksempel store datainnbrudd der datamengden er enorm.

Riksadvokaten har uttalt at ønsket om en høyere oppklaringsprosent aldri skal gå på bekostning av påtalemyndighetens vurdering av beviskravet.[130] Uttalelsen må forstås som at beviskravet i utgangspunktet er absolutt når påtalemyndigheten vurderer hvorvidt tiltale skal tas ut. Det oppstår her en interessekonflikt om hvorvidt beviset – som for oppgavens del antas å være et sentralt bevis – skal tillegges betydelig vekt, selv om man ikke kan svare på nødvendige spørsmål om beviset. Interessekonflikten åpenbarer en svakhet ved kontradiksjon rundt beviset som sannhetsfremmende mekanisme i det foreliggende tilfellet. Tradisjonell spørsmålstilling vil i et slikt scenario ikke skape klarhet rundt beviset, og hensynet til et materielt riktig resultat kommer i klem mellom et effektivt rettsvesen og det strenge beviskravet i strafferetten. Merk at i noen tilfeller er bruken av KI-verktøy med base i maskinlæring den eneste måten saken kan ha håp om å oppklares, eksempelvis ved større cyberangrep. Å akseptere en større usikkerhet rundt beviskravets oppfyllelse – herunder kravet til robusthet – vil gå på bekostning av tilliten til prosessen, mens et dysfunksjonelt rettsvesen vil gjøre det samme. Det kan på bakgrunn av dette argumenteres for at det strenge beviskravet i flere tilfeller kan hindre politiet fra å benytte slike systemer når det føres bevis i retten, jf. Riksadvokatens uttalelse om at beviskravet aldri skal senkes med det formål å oppnå en høyere oppklaringsprosent.[131] Dette resultatet vil være en naturlig konsekvens av at rettens aktører ikke vet hvor pålitelig KI-systemet er. Den åpenbare løsningen vil i møte med KI-problematikken være å gjennomgå og teste systemet slik at påliteligheten kan kvantifiseres. Gode grunner taler derfor for at problematikken må løses på et mer fundamentalt plan enn fra

---

[129] For sondringen rundt sannsynlighetskrav og robusthetskrav: Løvlie, A. *Rettslige faktabegreper*. Side 335-337.

[130] Riksadvokatens rundskriv nr. 3/2018. Revidert 21.02.2019. *Kvalitetskrav til straffesaksbehandlingen i politiet og ved statsadvokatembete mv. (kvalitetsrundskrivet).* Side 15.

[131] Ibid.

sak til sak. Omfattende testing utenfor den enkelte prosess vil utvilsomt være ressurskrevende, men vil igjen resultere i en langt mer effektiv straffeprosess i saker der KI-bevis føres, sammenlignet med spørsmålsrunder om programmets treningsdata, treningsmodeller, innkodede mål, etc.

Kontradiksjons som sannhetsfremmendemekanisme kan dermed ikke sies å ivareta de grunnleggende hensynene i straffeprosessen i møte med KI-problematikken. Det kan tenkes at kontradiksjon vil kunne fungere som rettssikkerhetsmekanisme for den tiltalte i enkelte tilfeller, men på generelt grunnlag vil mekanismen ikke systematisk ivareta rettssikkerheten til borgere. Politiet og påtalemyndigheten vil ofte havne i en situasjon der de ikke kan svare på kontrollspørsmål om programmets oppbygning og treningsdata. Det er samtidig uheldig ut ifra det grunnleggende hensynet til tillit at det skal være tiltaltes jobb å falsifisere politiets digitale metodebruk. Effektivitetshensyn taler for at problemet skal adresseres på et mer fundamentalt plan enn fra sak til sak. En lite effektiv prosess grunnet omfattende utspørring om treningsdata og oppbygning vil også svekke tilliten til rettsvesenet. Den alternative løsningen vil være å legge så liten vekt på KI-beviset – grunnet manglende tilfredsstillende svar – at det går ut over politiets evne til å bekjempe kriminalitet, noe som kan føre til uriktige frifinnelser.

Kontradiksjon som sannhetsfremmende mekanisme kan på bakgrunn dette ikke sies å få bukt med KI-problematikken. Det går nå over til neste rettssikkerhetsgaranti, påtalemyndighetens ansvar for et forsvarlig avgjørelsesgrunnlag.

3.2.4 Politiet og påtalemyndighetens ansvar for et forsvarlig avgjørelsesgrunnlag

Påtalemyndigheten har et ansvar for å sørge for et forsvarlig avgjørelsesgrunnlag og plikten er strengest når det tas ut tiltalte i de alvorligste straffesakene. Det som skal avgjøres i denne sammenheng er spørsmålet om tiltale. Når avgjørelsesgrunnlaget kan sies å være forsvarlig varierer fra sak til sak, og fra sakstype til sakstype.[132] Det må blant annet tas hensyn til utredningskravet, objektivitetskravet, kontradiksjonsretten, effektivitetshensyn og notoritetskrav. Kravet om forsvarlighet i straffesaker har Grunnlovs rang etter Grl. § 95.[133] I denne sammenheng rettes ansvaret mot bevisbildet i saken og forholdene rundt.

[132] Selv om det eksiterer et likhetsprinsipp i straffeprosessen vil hver enkelt sak være forskjellig.
[133] Jon Petter Rui, *Straffeprosessen i perspektiv.* Punkt 5.1.

For at politiet og påtalemyndighetens ansvar for et forsvarlig avgjørelsesgrunnlag kan sies å utgjøre en adekvat rettssikkerhetsgaranti i møte med utfordringene som kan følge med bevis produsert av KI-systemer, må politiet og påtalemyndighetens rutiner veie opp for den usikkerhet som eksisterer. Riksadvokaten har i «kvalitetsrundskrivet» uttalt at for å utferdige tiltale må påtalemyndigheten forholde seg til samme strenge beviskrav som domstolene.[134] En antitetisk tolkning av beviskravet i straffesaker taler for at politiet og påtalemyndigheten i utgangspunktet bare skal ta hensyn til «rimelig tvil».

Som tidligere nevnt er dagens politipraksis ikke tilrettelagt for at påtalemyndigheten får tilstrekkelig informasjon gjennom de rutiner som allerede eksisterer. Det er derfor vanskelig for påtalemyndigheten å overholde kravet om et forsvarlig avgjørelsesgrunnlag slik situasjonen er i dag. Det ble også slått fast ovenfor at dette i utgangspunktet ikke er tilstrekkelig da usikkerheten er for stor til å neglisjeres sett opp mot det strenge beviskravet. Problemet skyldes langt på vei at påtalemyndigheten ikke har positiv kunnskap om problematikken. Det som skal drøftes er hvor langt kravet til sakens opplysning strekker seg etter plikten til et forsvarlig avgjørelsesgrunnlag i hver enkelt sak der KI-bevis er involvert, og om plikten er en betryggende rettssikkerhetsgaranti.

Det er påtalemyndigheten som har ansvaret for etterforskningen, jf. påtaleinstruksen § 7-5 annet ledd[135], mens politiet utfører den etter strpl. § 225.[136] Det rettslige grunnlaget for politiets og påtalemyndighetens ansvar for å sørge for et forsvarlig saksgrunnlag framkommer av vilkårene for tiltalebeslutningen i strpl. § 249. Etter ordlyden skal saken være «tilstrekkelig forberedt» før avgjørelse om tiltale tas. Hvordan begrepet skal forstås er nærmere angitt i strpl. § 226 i kapittelet om etterforskningen.

Det fremkommer i § 226 at formålet med etterforskningen er å skaffe «skaffe til veie de nødvendige opplysningene for å avgjøre spørsmålet om tiltale». Politiet skal avdekke og gjennomgå bevis for å avgjøre om det har skjedd noe straffbart og eventuelt hvem som har begått lovbruddet. Etter § 226 tredje ledd skal etterforskningen søke å klarlegge de opplysninger som taler mot han, men også de som taler for han. Regelen gir uttrykk for

---

[134] Riksadvokatens rundskriv. *Kvalitetskrav til straffesaksbehandlingen i politiet og ved statsadvokatembete mv. (kvalitetsrundskrivet).* Side 15. Påtaleansvarlig må samtidig være overbevis om siktedes straffeskyld og at den kan føre fullgode, og lovlige, beviser for dette i retten.

[135] Forskrift 28. juni 1985 nr. 1679 om ordningen av påtalemyndigheten (Påtaleinstruksen).

[136] NOU 2016: 24 om ny straffeprosesslov side 301: «Til tross for uklar ordlyd er det ikke tvilsomt at det etter gjeldende rett er påtalemyndigheten i politiet som har det formelle og overordnede ansvar for etterforskingen i straffesaker, jf. også påtaleinstruksen § 7-5.»

objektivitetsplikten som også er lovfestet i stprl. § 55 siste ledd. Av objektivitetsplikten følger det blant annet at politi- og påtalemyndigheten har en aktiv plikt til å forfølge hypoteser der siktede er uskyldig.

I den forbindelse samler politiet bevis og spor som kan kaste lys over spørsmålet om tiltale. Politiet har en plikt til å behandle bevisene på en forsvarlig måte slik at de ivaretar sin integritet, samt at de ikke skades, forurenses, etc. Bevisene må være innhentet, analysert og behandlet på en rettferdig måte.[137] Regelen gjelder også for «kriminalteknisk»-arbeid.[138] [139] Den påtaleansvarlige har ansvaret for å vekte bevisene i saken på en forsvarlig måte når hun vurderer om saken er «tilstrekkelig opplyst for å avgjøre spørsmålet om tiltalte.»

Det forutsettes at den påtaleansvarlige vet at beviset er produsert av et KI-system og er kjent med utfordringene som kan følge med slike systemer. Den påtaleansvarlige vet derimot ikke hvor pålitelig det aktuelle programmet er. En naturlig følge vil da være at den påtaleansvarlige må få klarhet i dette før det legges avgjørende vekt på det sentrale beviset.

Spørsmålet blir da hvor langt den enkelte påtaleansvarlig må strekke seg i en sak med KI-bevis før plikten til å sørge for et forsvarlig avgjørelsesgrunnlag kan sies å være oppfylt. Den påtaleansvarlige må forholde seg til det system som er tilrettelagt for å etterleve dette kravet, samtidig som hun forholder seg til de faktiske forholdene. Det kan eksempelvis ikke kreves at den påtaleansvarlige må foreta noe som er umulig eller systemet ikke tillater. Det antas igjen at beviset må anses som et fellende bevis for at drøftelsen skal ha mest for seg.

Siden den påtaleansvarlige vet at det er et KI-system som har behandlet beviset kan det argumenteres for at hun burde prøve å skaffe seg en forståelse av hvor pålitelig beviset er, jf. kravet om notoriet rundt bevisinnhenting og håndtering av tekniske bevis.[140] Organiseringsstrukturen i Norge gir noe bedre adgang til dette ettersom politiet og påtalemyndigheten samarbeider tett sammenlignet med andre land. Den påtaleansvarlige har dermed en lettere tilgang til politiets dataetterforskere, som i dagens system – ved siden av de som har kjøpt inn programmet – er de som er best egnet til å svare på spørsmål om de digitale

---

[137] ECHR, Van Mechelen and others v. The Nederlands, 30 October 1997 (Saksnummer 55/1996/674/861-864) § 50.
[138] ECHR, Gül v. Turkey, 14 December 2000, (Saksnummer 22676/93) § 89.
[139] Med «kriminalteknisk»-arbeid siktes det til «forensics»-arbeid. Begrepet omfatter etterforskningsarbeid som ofte krever spisskompetanse for å utføre. Arbeidet har ofte tettere koblinger til vitenskap enn mer tradisjonelt etterforskningsarbeid som f.eks. avhør.
[140] Riksadvokatens rundskriv. *Kvalitetskrav til straffesaksbehandlingen i politiet og ved statsadvokatembete mv. (kvalitetsrundskrivet).* Side 28. Punkt 4.11.2. Om krav til notoritet i bevisbehandling.

verktøyene. Imidlertid kan det i dagens system være tilfelle at programmet som har behandlet beviset ikke er testet eller at man ikke har innsyn i programmets oppbygning.

Undersøkelsene den påtaleansvarlige gjør kan resultere i at «out-put»-et dobbeltsjekkes eller at programmet testes med andre «in-put»-data, for å se hvordan programmet reagerer. Rutinen må sies å være sannhetsfremmende da den er med på å minske usikkerheten. På den andre siden vil det å kjøre programmet flere ganger bare sikre mot de mest åpenbare feilene i systemet. Mer underliggende feil vil vanskelig kunne oppdages på denne måten da det krever mer omfattende testing. Det kan samtidig ikke forventes at påtalejuristen skal iverksette omfattende testing for å kunne sies å ha oppfylt sin plikt til å sørge for et forsvarlig avgjørelsesgrunnlag. En omfattende gjennomgang av koder og treningsdata (dersom disse mot formidling er tilgjengelig) vil raskt komme i konflikt med den siktedes rett til å få sakens avgjort innen «rimelig tid», jf. Grl. § 95 og EMK art. 6.

Utenfor den enkelte sak er det per i dag vist liten vilje til å iverksette slik omfattende testing. Det kan argumenteres for at digitaletterforskning i dag er underfinansiert, samtidig som digitale testverktøy konkurrer i samme budsjett som skuddsikre vester og politibiler. Fagmiljøene er samtidig lavt organisert og sliter dermed å få budskapet sitt gjennom innad i politiet.[141]

Det vil da være faktiske- og systemutfordringer, og ikke rettslige, som hindrer en solid utredning av KI-programmenes pålitelighet. Det ansvaret som kan legges på hver enkelt påtaleansvarlig ut fra utredningskravet begrenses dermed av de faktiske forholdene. Det kan dermed neppe sies at ansvaret for å sørge for et forsvarlig avgjørelsesgrunnlag strekker seg lenger enn at den påtaleansvarlige meddeler de andre aktørene i rettssystemet at beviset er produsert av et KI-system. Den som treffer et positivt påtalevedtak om tiltale vil selvsagt ha et ansvar for å veie beviset på en forsvarlig måte opp mot den usikkerhet KI-systemet må antas å operere under.

For å unngå at det legges for mye vekt på det KI-behandlede beviset vil det mest åpenbare svaret vil være at den påtaleansvarlige bygger saken slik at beviset som er behandlet av KI-systemet, kun inngår i en beviskjede, hvorav flere andre bevis støtter opp om bevistemaet. Den påtaleansvarlige trenger da ikke å lene seg på KI-beviset i like stor grad. Imidlertid vil fordelene med å benytte KI-verktøy i utgangspunktet raskt bli spist opp av ulempene,

[141] Dataetterforskningsmiljøet er ikke spesielt høyt plassert i hierarkiet i politiet, noe som vil være en faktor når problemet skal adresseres.

ettersom verktøyet ble tatt i bruk for å kunne levere fullgode beviser. Manglende testing vil bety at man må forholde seg til dataetterforskningsverktøy med base i KI med en «antatt error rate».[142] I den forbindelse vil den primære funksjonen til KI-systemer bli å være et ledd i etterforskning der funnene må suppleres med andre bevis gjennom ytterligere etterforskning. Effektivitetshensyn taler helt klart mot en slik løsning da bevisanalyseverktøy utgjør en viktig del av bevisbehandling. Å iverksette etterforskingsskritt for å sikre bevismateriale ved siden av bevis som er produsert av kunstig intelligens, kan føre til at saker blåses uforholdsmessig mye opp. Resultatet vil da bli at hver enkelt sak krever mer ressurser og færre saker løses.

En bedre løsning vil helt klar være å redusere den usikkerhet man allerede vet eksisterer og opprettholde bevisverdien mest mulig. Dette kan – som allerede nevnt - gjøres gjennom systematisk testing av programvaren. Gode grunner taler derfor for at systemet må endres for at bevisene skal kunne tjene sitt formål på best mulig måte.

Påtalemyndighetens ansvar for å sørge for et forsvarlig avgjørelsesgrunnlag kommer til kort i møte med KI-problematikken. Når problematikken blir kjent og andelen KI-verktøy øker, vil den påtalejuristen som har ansvaret for hver enkelt sak finne det vanskelig å forholde seg til KI-bevisene i dagens system. Usikkerheten som må legges til grunn når programmet ikke er testet, kan gjøre at bevisets bevisverdi må senkes uforholdsmessig mye i bevisvurderingen, jf. «at enhver rimelig tvil skal komme den siktede til gode».

3.2.5 Rettens ansvar for et forsvarlig avgjørelsesgrunnlag

Retten har et selvstendig ansvar for å sørge for sakens opplysning etter strpl. § 294. For tekniske beviser som krever en særkompetanse, vil sakkyndigordningen være rettssikkerhetsmekanismen som skal skape klarhet i bevissituasjonen. For at rettens plikt etter strpl. § 294 kan sies å adresse KI-problematikken på en adekvat måte må den – i likhet med de øvrige mekanismene – ivareta de grunnleggende hensynene i straffeprosessloven.

Rettens ansvar for å sørge for et forsvarlig avgjørelsesgrunnlag etter strp. § 294 er i utgangspunktet sekundær da det er partene som i utgangspunktet har ansvaret for sakens opplysning, og da særlig påtalemyndigheten som i utgangspunktet har bevisbyrden.[143]

---

[142] «Antatt error rate» vil i oppgavens øyemed sikte til det faktum at politiet vet at det eksiterer en ikke ubetydelig «error rate», men at de ikke vet hva den ligger på. De må dermed gjøre en antagelse, og for at antagelsen skal være rettmessig må «rimelig tvil om faktiske forhold komme den siktede til gode».
[143] Unntak finnes i lov 18 juni 1965 nr. 4 om vegtrafikk (vegtrafikkloven) § 22 femte ledd.

Etter strpl. § 294 skal retten «våke over at saken blir fullstendig opplyst.» En naturlig språklig forståelse av ordlyden kan tyde på at saken ikke kan avgjøres lovlig før alle stener i saken er snudd. En slik forståelse kan imidlertid ikke legges til grunn da ordlyden må modereres mot effektivitets- og ressurshensyn. Ordlyden «fullstendig» kan således ikke tas på ordet, men er snarere en påminnelse om at saken må være tilstrekkelig opplyst for å danne et forsvarlig avgjørelsesgrunnlag.[144] For at det skal foreligge et forsvarlig avgjørelsesgrunnlag som kan gi grunnlag for en strafferettslig dom må avgjørelsesgrunnlaget være robust.[145] Robusthetskriteriet må sees i sammenheng med beviskravet i strafferetten. Med dette menes at retten med stor sikkerhet skal være overbevist om at bevisgrunnlaget er korrekt, og at bevismaterialet er robust. Eksempelvis vil to identiske vitneutsagn fra to uavhengige vitner uten egeninteresse i saken være et mer robust enn ett vitneutsagn. En sak kan vanskelig sies å være uten noen form for rimelig tvil hvis bevisgrunnlaget ikke er tilstrekkelig robust. For eksempel kan bevisbildet i en straffesak peke mot straffeskyld, men likevel være så sårbart for ny informasjon at det ikke er naturlig å konkludere med at beviskravet er nådd.[146] Det kan da sies at bevisgrunnlaget ikke er tilstrekkelig robust. Et praktisk eksempel er der etterforskerne ikke har undersøkt tiltaltes alibi. Den alternative forklaringen har potensial til å velte hele saken politiet har mot den tiltalte, og må således undersøkes.

Robusthetskravet er en side av beviskravet som kun er ment å ivareta den tiltaltes interesser, mens domstolens generelle utredningsplikt etter § 294 skal ivareta sakens opplysning i begge retninger, herunder kartlegge den materielle sannhet. Således er kravet om robusthet bare begrunnet i én retning, nemlig å unngå uriktige domfellelser, mens utredningsplikten gagner begge parter. Selv om retten normalt kan basere seg på det bevismateriale aktor legger frem, kan saken likevel ta en slik vending at retten plikter å sørge for supplerende bevisføring før det avises dom, jf. Rt. 2013 s. 905 avsnitt 39. I Rt. 2008 s. 605 avsnitt 14 tolker Høyesterett innholdet i strpl. § 294 slik:

«Rekkevidden av rettens plikt etter § 294 beror på sakens omstendigheter, med det generelle bevisbilde som utgangspunkt for vurderingen. Et sentralt moment er hvor alvorlig saken er. I tillegg vil det kunne ha betydning om opplysningene gjelder et viktig punkt i saken, om de

[144] Øyen, Ø. *Straffeprosess*. Side 349.
[145] Løvlie, A. *Rettslige faktabegreper.* Side. 166.
[146] Op.cit. Side 165.

antas å være til tiltaltes gunst, og om de kan fremskaffes ved rimelig innsats av tid, penger og andre ressurser i forhold til hva man må regne med kan oppnås.»

Selv om utgangspunktet er at rettens plikt til å undersøke forhold langt på vei styres av partenes anførsler viser kjennelsen fra 2008 at retten gjennom strpl. § 294 har en aktiv plikt til å sørge for sakens opplysning.[147] Det vil i stor grad være dette robusthetskravet retten må forholde seg til når den skal idømme straff på bakgrunn av et bevis som er produsert av et KI-system, jf. de usikkerhetsmomentene og mulige feilkilder som det er redegjort for i kapittel 2. For spørsmål rundt KI-bevis vil oppnevnelsen av sakkyndige være den naturlige veien å gå for å få klarhet i et bevis med base i et KI-system.

3.2.5.1 Sakkyndighetsordningen

Den rettslig oppnevnte sakkyndige oppnevnes av retten med hjemmel i strpl. § 138. Formålet med å involvere sakkyndige er få klarhet i bevistemaer der det kreves spisskompetanse. Utgangspunktet er at det oppnevnes én sakkyndig, jf. strpl. § 139. Det som må undersøkes er hvor godt den/de sakkyndige kan bidra til saken opplysning på generelt grunnlag.

For det tilfelle at programmet som har behandlet beviset er laget av politiet, vil den ansatte som utviklet programmet være den som er nærmest til å oppklare usikkerheten rundt KI-systemet. Imidlertid vil det – som nevnt – ofte være de mest sofistikerte programmene som er gjenstand for den særskilte KI-problematikken, herunder «svart boks»-problemet. Uheldigvis er det også slike systemer politiet kjøper fra leverandører som ofte gjemmer programmets kapasiteter bak «lukket» kildekode.

Risikoen er at de sakkyndiges vitneprov blir mer generell enn man kunne ønske. Et vitneprov fra en sakkyndig på generell basis vil ikke si noe konkret om den aktuelle programvaren som har blitt benyttet, men snarere være en utredning av problematikken rundt beslektede systemer med tilsvarende kapasiteter. Senere rettsutvikling har også sett en økning i adgangen til private sakkyndige etter prinsippet om fri bevisføring.[148] Det er derfor plausibelt at flere sakkyndige kan ende med å forklare seg for retten. I likhet med mange andre fagdisipliner er det uenighet knyttet til kapasiteter, styrker og svakheter ved teorier, teknologier og teknikker, og digital kriminalteknikk er intet unntak. Det vil derfor være en relativt lett øvelse å finne noen som argumenterer i egen favør enten man er påtalemyndighet eller tiltalt. Adgangen til fri bevisføring med påfølgende snevre avskjæringsregler gjør at rettssystemer, herunder

[147] Ørnulf Øyen på side 158 i Hedlund, mfl. *Bevis i straffesaker: utvalgte emner*.
[148] Nils Erik Lie på sidene 59-63 i Hedlund, mfl. *Bevis i straffesaker: utvalgte emner*.

bevisbedømmerne, står i fare for å havne i midten av en meningsutveksling mellom to «eksperter». Fenomenet omtales som «balle of experts» og er noe rettsstaten Norge har forsøkt å unngå.[149] Slike tilfeller oppstår ikke rent sjeldent i rettssystemer med tradisjon for private ekspertvitner. Eksempelvis var dette tilfellet i Loomis vs Wisconsin der tvistespørsmålet var påliteligheten av KI-systemet.[150] Når norske dommere skal avgjøre spørsmålet vil de prosessuelle omstendighetene sannsynligvis være de samme; en leverandør som ikke vil utgi programvaren grunnet forretningshemmeligheter og motstridende ekspertuttalelser. En slik situasjon vil stille høye krav til dommerne i saken. Det kan derfor vanskelig sies at sakkyndigordningen vil bidra til noen umiddelbar løsning av problematikken.

For den tiltaltes vedkommende vil imidlertid involvering av sakkyndige være gunstig. At den tiltalte får dratt prosessen inn i et spor med sakkyndiguttalelser vil bidra til å sette lyst på problemet. Et slikt utgangspunkt gir en langt bedre forutsetning for å unngå at feilaktige bevis legges til grunn enn om utfordringene ikke problematiseres. På den andre siden vil det være kostbart både for den tiltalte og samfunnet om det skulle oppnevnes sakkyndige hver gang det kommer en innsigelse mot et slik digitalt etterforskningsverktøy. For de tilfeller at det er adgang til koder og treningsdata vil effektivitet- og kostandshensyn med styrke tale mot at problemet løses på en «sak til sak» basis. Et annet hensyn som taler imot er hensynet til en tillitsvekkende prosess, der man ved å overlate til retten å avgjøre spørsmålet om programmets pålitelighet, skyver ansvaret for kontrollen med KI-systemet over på retten og den tiltaltes forsvarer.

Med de virkemidler dommerne har til rådighet vil de ha utfordringer med å nærme seg den «materielle sannhet», slik det kommer til uttrykk i «det materielle sannhets prinsipp». Ser en til andre spesialfelter – eksempelvis medisin – kan de sakkyndige legge fram et fortrolig resonnement der vurderingen og erfaringssetningene kontrolleres av retten, samt av den rettsmedisinske kommisjonen. For det tilfellet at programvaren – som baser seg på en avansert maskinlæringsmodell – har fremskaffet et bevis, vil de sakkyndige ikke kunne følge og verifisere resonnementet. Under slike forhold vil den sakkyndiges vurderinger miste mye

[149] NOU 2001: 12 Rettsmedisinsk sakkyndighet i straffesaker. Side 137.
[150] Loomis vs Wisconsin avsnitt 46.

av sin tyngde ettersom den sakkyndige selv er overlatt til kvalifiserte gjetninger, og ikke reelle forklaringer (årsak, virkning).[151]

Sakkyndigordningen vil dermed ikke bidra til å løse problemet i nevneverdig stor grad, og følgelig vil kostnadene ordningen fører med seg ikke være forholdsmessige. Effektivitetshensyn taler imot at det skal være opp til sakkyndigordningen – grunnet det begrensede bidraget – å skape klarhet i problematikken rundt bevis produsert av kunstig intelligens. I likhet med de øvrige mekanismene vil rettens ansvar for sakens opplysning ikke kunne møte problematikken på en tilfredsstillende måte. Føring av denne type bevis uten de nødvendige kontrollmekanismene vil kunne gå på bekostning av tilliten til rettsvesenet. Retten lener seg normalt på de sakkyndiges evne til å skape klarhet i bevismassen når det gjelder beviser tekniske beviser. I denne sammenheng blir imidlertid retten i stor grad overlat til seg selv grunnet det begrensede bidraget fra sakkyndige.

3.3 Oppsummering og konklusjon

Innledningsvis ble det vist at det eksisterer flere utfordringer knyttet til introduksjonen av kunstig intelligens i bevisbehandling. En av de største utfordringene knyttet seg til usikkerhet og feilkilder ved slike systemer. Samtidig utfordres de tradisjonelle kravene til bevisets notoritet og dokumentasjon gjennom «svart boks»-problemet som ofte følger med mer sofistikerte maskinlæringsmodeller. Utfordringene sett i sammenheng taler for en robust kontroll med de systemene som benyttes, og hvordan de er bygget opp. Leverandørene av dataetterforskningsverktøyene er på samme tid tilbakeholdne med å dele informasjon om programmene, noe som legger ansvaret for programmets pålitelighet over på politiet. På samme tid bør man erkjenne at problemet er for stort til å neglisjeres, jf. usikkerheten opp mot det strenge beviskravet i strafferetten. Det er derfor nødvendig å innføre kontrollmekanismer som testing.

Det er imidlertid ikke rutine i norsk politi i dag å teste programmene eller foreta andre former for usikkerhetsminimering. Når programmene først testes skjer dette på en altfor liten skala sammenlignet med hva som er nødvendig.

Resultatet av manglende empiriske data gjør at bevisbedømmerne må benytte seg av antagelser når de vurderer beviset. Antagelsene har en tendens til å trekkes ut ifra en analogi

[151] Kolflaath, E. 1. utg. 2013. *Bevisbedømmelse i praksis*. Bergen: Fagbokforlaget Vigmostad & Bjerke AS. Side 163.

fra digitale systemer som ikke baserer seg på avanserte maskinlæringsmodeller. Som nevnt opereres det generelt med en større usikkerhet rundt avanserte maskinlæringsmodeller. På bakgrunn av dette resonnementet er det viktig at denne informasjonen når rettens aktører, og da særlig dommerne. Imidlertid er det ingen praksis i norsk politi i dag å opplyse om at KI-systemer har vært involvert i bevisbehandlingen, noe som kommer i konflikt med tiltaltes rettigheter. Manglende opplysning om forholdene gjør at rettssikkerhetsmekanismene som skal ivareta tiltaltes rettigheter ikke kommer i spill. Det står klart for meg at en endring i praksis er nødvendig på dette punktet.

Norsk straffeprosess har mekanismer som er ment å skape klarhet i og rundt bevis som føres for retten. Gjennomgangen av rettssikkerhetsmekanismene har imidlertid vist at de ikke er adekvate til å ivareta de underliggende hensynene som norsk straffeprosess bygger på, i møte med bevis produsert av KI-systemer. Konsekvensen vil være at beviset som er produsert av slike systemer ikke kan gis den bevisverdi som trengs for å opprettholde en effektiv kriminalitetsbekjempelse i et stadig skiftende kriminalitetsbilde. Det oppstår dermed en interessekonflikt mellom robusthetskravet og ønsket om en effektiv kriminalitetsbekjempelse.

Balansegangen der alle hensynene ivaretas på et akseptabelt nivå er vanskelig å foreta i dagens system, og hensynene vil kunne gå uforholdsmessig mye ut over hverandre. Mye tyder derfor på at dagens rettssystem, de lege lata, kommer til kort i møte med KI-problematikken for bevisbehandling i straffesaker. Hypotesen i oppgaven er derfor bekreftet. Analysen har gjort det tydelig at utfordringene introduksjonen av kunstig intelligens i bevisbehandling utgjør må løses på et mer fundamentalt plan enn dagens «fra sak til sak»-system. Det skal i neste kapittel skal drøftes hvordan utfordringene kan løses.

4 Del 2 - De lege ferenda

Dette kapittelet vil se på hvilke potensielle tiltak som kan – og bør – innføres for å møte utfordringene KI i bevisbehandling representer. Oppgaven vil i så måte være en rettspolitikk drøftelse der målet er å sikte mot en løsning som er rimelig effektiv og tillitvekkende, samtidig som tiltaltes rettssikkerhet ivaretas. Dette gjøres gjennom å vurdere de faktiske forholdene opp mot de grunnleggende hensynene i strafferettspleien.

4.1 Hva bør gjøres?

Det er hensiktsmessig å ta utgangspunkt i hovedproblemet, nemlig den usikkerhet som kan følge med KI-systemer. For at bevis som er behandlet av KI-systemer skal kunne tjene sitt formål i den utstrekning som er tiltenkt, må usikkerheten kvantifiseres, jf. kapittel 2. Dersom

usikkerheten ikke kvantifiseres vil en måtte legge inn et forbehold når beviset presenteres eller gjøre en antagelse. Antagelsen må i så måte gjøres slik at «rimelig tvil» om faktiske forhold kommer den tiltalte til gode. En slik antagelse vil svekke beviskraften uforholdsmessig mye dersom den egentlige påliteligheten av programmet er høy. Det mest hensiktsmessige vil dermed være å undersøke programvaren. Det vil også være hensiktsmessig å teste programvaren før bruk, snarere enn fra sak til sak.

Resonnementet finner støtte i EUs utredning om kunstig intelligens implementering i samfunnet. Kommisjonen og EUs ekspertkomite har tatt til orde for at problemene ved implementeringen av KI-systemer i samfunnet burde løses på nasjonalt nivå gjennom kontrollmekanismer.[152] Det er særlig uttrykt bekymring for «high-risk»-applikasjoner. Slike «high-risk»-applikasjoner kjennetegnes ved at de utfører oppgaver på et område der konsekvensene kan gjøre stor skade eller true fundamentale rettigheter.[153] Det er ikke tvilsomt at kriminalitetsbekjempelse og straffeforfølgning er et slik område.[154] EU Kommisjonen har i sin white paper spesifikt utpekt digitale etterforskningsverktøy (dvs. verktøy for digital behandling av bevis) som «high-risk». EU kommisjonen har også kategorisert KI-systemer som identifiserer personer ved hjelp av biometriske identifikasjoner på avstand som alltid «high-risk»-applikasjoner. Det omfatter blant annet fingeravtrykk, ansiktsgjenkjenning, iris-scan, vaskulære mønstre, etc. Disse typene identifikasjonsteknikker kan tenkes å bli benyttet av norsk politi i dag eller i fremtiden.[155] Et praktisk eksempel kan være norsk politis nylige bruk av droner.[156] Ved en alvorlig hendelse kan det tenkes at politiet rykker ut med droner og videomaterialet fra hendelsen kan senere brukes i retten. Det kan da tenkes at ansiktsgjenkjenning blir brukt til å identifisere en person som løper fra stedet. Ut ifra EU-kommisjonens forståelse er det avgjørende at KI-systemene kontrolleres gjennom hele sin livssyklus, ikke bare ved erverv.[157] En løsning er således avhengig av tett og jevnlig kontroll av systemene.

---

[152] Ekspertgruppens arbeid har munnet ut i white papaer (2020) etter at gruppen først la frem sine retningslinjer i mars/arpil 2019.
[153] *White Paper on Artificial Intelligence: a European approach to excellence and trust.* Side 17.
[154] *White Paper on Artificial Intelligence: a European approach to excellence and trust*. Inntatt i fotnote 50 på side 17.
[155] Under forutsetninger at de følger Article 9 i General Data Protection Regulation (GDPR). Forordningen er inntatt i lov 15 juni 2018 nr. 38 om behandling av personopplysninger (personopplysningloven) § 1. og The European Parliament and of the Council. Directive 2016/680 (27.04.2016). *On Law Enforcement Directive.* Article 10.
[156] Aftenpostens artikkel: «Oslo-politiet tester utrykning til oppdrag med droner fra fredag» Tilgjengelig på: https://www.aftenposten.no/osloby/i/LAwEzp/oslo-politiet-tester-utrykning-til-oppdrag-med-droner-fra-fredag (sist lest 29.05.20)
[157] *White Paper on Artificial Intelligence: a European approach to excellence and trust*. Side 21 og 22.

En strengere kontroll av teknologien har også støtte i litteraturen.[158] Bevisteoretikeren Nils Erik Lie har tatt til orde for at framtidens dommere vil ha utfordringer i møte med teknologiske nyvinninger som politiet benytter i etterforskningen. I 2015 uttalte han at politiet – gjennom mer avansert teknologi – vil sette sitt preg på bevisbildet i straffesaker, og han uttrykker bekymring for at fremtidens dommere kan trenge gjennom den tekniske materien. Lies uttalelser må forstås som en frarådelse mot at det overlates til dommere å føre den strenge kontrollen med den avanserte teknologien. Lie kommer imidlertid ikke med noe forslag til hvordan utfordringene kan løses, men snarere at det kan bli et problem.

Slik jeg ser det har det formodningen mot seg å overlate sakens opplysning til dommerne etter strpl. § 294, herunder sakkyndigordningen. En løsning burde heller søkes utenfor hver enkelt sak, og da slik EU-kommisjonen oppfordrer til, på nasjonalt nivå. Ved å møte problemet forut for en potensiell rettslig tvist kan en unngå at mange av tvistene oppstår i utgangspunktet. En forsvarer som vet at KI-systemet som har behandlet beviset er 98 % pålitelig vil ha mindre interesse av å angripe beviset enn hvis KI-systemet må antas å ha en pålitelighet mellom 80-99 %. Å ta problemet forut for straffesakene vil således ha støtte i effektivitet- og rettferdighetshensyn. Rettferdighetshensynet ivaretas gjennom antagelsen at mer informasjon om programmets pålitelighet danner grunnlag for en bedre bevisbedømmelse.

Hensynene som taler for en slik overordnet kontrollordning er først og fremst hensynet til et materielt riktig resultat, effektivitet i straffeprosessen og tilliten til rettssystemet. Hvorvidt en slik ordning vil være kostnadseffektiv og i så fall hvor mye, vil avhenge av omfanget av kontrollordningen.

Når det tas til orde for at en ny kontrollordning burde innføres vil det være av interesse å se til den eneste ekspertorienterte kontrollordningen som allerede eksisterer i straffeprosessen.

4.2 Utgangspunkt i den rettsmedisinske kommisjon

I Norge er det kun den rettsmedisinske kommisjonen som har en lovfestet status som ekspertkommisjon i straffeprosessen etter strpl. § 146. Kommisjonen ble opprettet i 1900 og har som hovedoppgave å kvalitetssikre sakkyndige uttalelser i tilknytning til rettsmedisinsk vitenskap og kunnskap.[159] Behovet for sakkyndighet i straffeprosessen var allerede en realitet når kommisjonen ble opprettet. Det Medicinske Fakultet ved Københavns Universitet avga

[158] Nils Erik Lie på side 67 i Hedlund, mfl. *Bevis i straffesaker: utvalgte emner*.
[159] Anders Løvlie på side 555 i Hedlund, mfl. *Bevis i straffesaker: utvalgte emner*.

allerede på 1600-tallet rettsmedisinske erklæringer i straffesaker.[160] Ut ifra behovet som meldte seg i 1900 ble den rettsmedisinske kommisjonen i Norge opprettet. En viktig bakgrunn for opprettelsen var å ivareta den tiltaltes rettssikkerhet.[161] Ordningen har gjennom årene utviklet seg, men kommisjonens hovedoppgaver er langt på vei de samme. For ca. 20 år siden opprettet stortinget et ekspertutvalg som skulle vurdere den rettsmedisinske kommisjonen. Utvalgets utredning ble inntatt i NOU 2001: 12 om rettsmedisinsk sakkyndighet i straffesaker.

I dag eksisterer ordningen for at rettens aktører skal gis et mest mulig riktig inntrykk av forholdene i saken.[162] Dommerne vil gjennom ekspertenes uttalelser ha den beste forutsetningen for å etterleve det materielle sannhets prinsipp, jf. at dommerne er eksperter på juss, men ikke har nødvendig faglig innsikt i rettsmedisin. Siden de rettsmedisinske undersøkelsene ofte utgjør et sentralt bevistema i straffesakene vil de sakkyndiges bidrag ofte være avgjørende.[163] Det har derfor vært særlig viktig å kunne stole på uttalelsene som blir gitt i straffesaker. Det er videre et mål å redusere forekomsten av kontradiktorisk ekspertise («battle of experts») i straffesaker.[164] Det er lett å dele utvalgets inntrykk her ettersom land som tradisjonelt har gjort private sakkyndige til en hovedregel har betalt en høy pris. Forskning viser for eksempel at ekspertuttalelser av lav kvalitet er en av hovedårsakene til justismord i USA, 27 % av tilfellene.[165] U.S Federal Rules of Evidence (FRE)", Rule 702 og «Daubert Standarden» gir en bred adgang til å føre «eksperter» og de vil i større grad oppleve fenomenet «battle of experts».[166] Et digitalt kontrollorgan vil i likhet med den rettsmedisinske kunne kontrollere de opplysningene som legges for retten. Et land som har fått på plass en slik ordning er Nederland.

Nederland opprettet 1. januar 2010 NRGD (The Netherlands Register of Court Experts). NRGD er en uavhengig gruppe eksperter som bistår aktørene i straffeprosessen ved henvendelse. At organet er uavhengig – i likhet med den rettsmedisinske kommisjonen – er egnet til å øke tilliten til organet. Den rettsmedisinske kommisjonen er i Norge helt eller delvis plassert under universitetene. Det virker derfor fornuftig – av hensyn til tilliten – å

---

[160] Op.cit. Side 540.
[161] NOU 2001: 12 Rettsmedisinsk sakkyndighet i straffesaker. Side 137.
[162] Systembetraktninger av Forskrift 13. februar 2018 nr. 240 om Den rettsmedisinske kommisjon og NOU: 2001: 12 side 137.
[163] NOU 2001: 12. Side 137.
[164] Ibid.
[165] Saks, M. J. og Koehler, J. J. *The Coming Paradigm Shift in Forensic Identification Science.* Publisert i *Science*, 5. august 2005. s. 892-894.
[166] U.S Federal Rules of Evidence (FRE)", Rule 702. «Testimony by Expert Witnesses». Tilgjengelig på https://www.law.cornell.edu/rules/fre/rule_702

plassere et digitalt kontrollorgan utenfor politiet. I NRGD føres det streng kontroll med hvem som får ta del i gruppen og hvilket faglig informasjonsgrunnlag vurderingene skal skje på. På NRGDs hjemmeside står det at gruppen *«guarantees and promotes the consistent quality of the contribution made by court experts to the legal process.»*[167] NRGD-sakkyndige bidrar både med «hands on» og «hands off» etterforskningsarbeid, samtidig som de fungerer som sakkyndige vitner i retten for både tiltalte, dommere og påtalemyndigheten. NRGD begrenser seg ikke til digital ekspertise, men tilbyr sakkyndighet innen medisin og psykologi, DNA-analyser, våpen og ammunisjon, narkotika og gift, skriftanalyse, etc. Opprettelsen av organet knytter sammen juss og vitenskap og har vist gode resultater. Opprettelsen av NRGD viser at en opprettelse av et lignende organ i Norge ikke er så radikalt som en kanskje skulle tro, og nå som behovet har meldt seg her til lands er det noe som burde undersøkes nærmere.

I utvalgets utredning av NOU 2001: 12 uttales det på side 136 at *«Det kan være vanskelig på prinsipielt grunnlag å begrunne særbehandlingen av det rettsmedisinske sakkyndighetsbevis i forhold til ethvert annet bevis i en straffesak.»* Og på samme side *«Etter utvalgets oppfatning kunne kanskje tilsvarende kontrollmuligheter ha vært ønskelig både på andre fagområder og utenfor strafferettens område, men det faller utenfor utvalgets mandat å vurdere dette nærmere.»*

Utvalget åpner dermed opp for at det – etter deres mening – kan være behov for kvalitetskontroll på andre fagområder enn det rettsmedisinske. Utvalget går så over til å vurdere om dagens rettsmedisinske kommisjonsordning ivaretar det overordnede hensynet i vurderingen, rettssikkerheten.[168] Utvalget vurderer så eksisterende og alternative kvalitetssikringsmekanismer, både overordnede og de som kan implementeres fra sak til sak. Vurderingene taler for at en eventuell kommisjon ikke nødvendigvis må følge i den rettsmedisinske kommisjonens fotspor hva gjelder kvalitetssikringsmekanismer, men snarere at mekanismene som innføres skal ha som overordnet mål å ivareta de involvertes rettssikkerhet. I vurderingen av om det er behov for opprettelsen av et slik organ vil det avgjørende være hvor stor risiko rettssystemet står overfor. Det er på det rene at digitale systemer kan inneholde feilkilder og at problemet er presserende, se blant annet følgende avhandlinger: «Om informationstekniskt bevis», «The Paradox of Automatisation» og «Fallacies when Evaluating Digital Evidence Among Prosecutors in the Norwegian Police

[167] The Netherlands Register of Court Experts (NRGD) Hjemmeside https://english.nrgd.nl/
[168] NOU 2001: 12. Rettsmedisinsk sakkyndighet i straffesaker. Side 138.

Service».[169] I tillegg – som vist i denne avhandlingen – utgjør KI-systemer en ytterligere risiko for feilkilder, samt at straffeprosessystemet ikke ivaretar grunnleggende hensyn på en tilfredsstillende måte. Samtidig antas det at andelen beviser med base i KI er ventet å øke. Fremveksten av «smart cities»[170] og 5G-nettverk vil antagelig mangedoble andelen informasjon samfunnet produserer, noe som vil aktualisere bruken av KI-verktøy ytterligere. Når mengden data øker for hver dag og KI-systemer implementeres i etterforskning og bevisbehandling, taler gode grunner for at det må føres en mer systematisk kontroll med bruken, jf. kapittel 2 og EU kommisjonens oppfattelse. Selv om Norge skulle velge å utsette diskusjonen vil et mer grenseoverskridende kriminalitetsbilde føre til at EU, EMD og kriminelle aktører påtvinger oss problematikken. Det er generelt en god idé å være med på diskusjonen rundt slike spørsmål, snarere enn å hente seg inn.

Den overordnede begrunnelsen for opprettelsen av den rettsmedisinske kommisjonen var at dommerne i straffesaker fant det vanskelig å forholde seg til tekniske spor og beviser. Tolkningen av disse sporene var særlig vanskelige og krevde en dyptgående kunnskap. På samme måte vil en kunne argumentere for at datavitenskap i dag krever en slik ekspertise for å kunne vurdere spor og beviser med en stor grad av sikkerhet. På samme tid som det innenfor medisin ikke vil være forsvarlig å la enhver medisiner opptre for retten som sakkyndig, vil det heller ikke være forsvarlig å la enhver person med utdanning innenfor data opptre som sakkyndig.[171] I likhet med opprettelsen av den rettsmedisinske kommisjonen truer også dagens digitale etterforskningssituasjon rettssikkerheten til borgerne. En forskjell er at utfordringene som følger med avanserte maskinlæringssystemer i all hovedsak ikke lar seg løse fra sak til sak. Det betyr imidlertid ikke at en sak til sak kontroll ikke er viktig. Det vil fortsatt være nødvendig å undersøke om KI-systemet er brukt innenfor sitt nisjeområde eller om bruken er gjort i henhold til aksepterte standarder eller instrukser. Det kan derfor argumenteres for at dagens situasjon trenger en aktiv kontroll av digitale etterforskningssystemer for å møte kravene til rettssikkerhet, herunder det materielle sannhet prinsipp og statens ansvar for en forsvarlig saksbehandling.

---

[169] Erlandsen, T-E. (2019). *Fallacies when Evaluating Digital Evidence Among Prosecutors in the Norwegian Police Service.* (Masteroppgave, NTNU, Norges teknisk-naturvitenskapelige universitet). Tilgjengelig på: https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2617771
Borhaug, T. S. (2019). *The Paradox of Automation in Digital Forensics.*
Ekfeldt, J. (2016). *Om informationstekniskt bevis.*
[170] IBM (International Business Machines) som er en av verdens største teknologibedrifter definerer smart cities slik: "one that makes optimal use of all the interconnected information available today to better understand and control its operations and optimize the use of limited resources".
[171] NOU 2001: 12 s. 137.

Samtidig som grunnleggende straffeprosessuelle hensyn taler for en tett kontroll med KI-systemene vil også politisk press før eller siden skyve oss i den retningen. Som nevnt arbeider EU med implementeringen av KI-systemer i Europa, samtidig som mulige fallgruver utredes. Det er også utarbeidet et utkast til en konvensjon om elektroniske beviser, der artikkel 4 er særlig interessant: Article 4 – Authentication of electronic evidence 1. The party seeking to introduce electronic evidence in any legal proceeding has the burden of proving it is what it claims to be.[172]

Det snakkes altså om å pålegge en positiv bevisbyrde for bevisets pålitelighet for elektroniske bevis. Påliteligheten det siktes til vil i norsk bevisteori best beskrives som bevisets «robusthet». Fra ca. år 2000 har diskusjonen om bevisets robusthet vokst fram i norsk bevisteori og tankesettet i artikkel 4 i utkastet er ikke ukjent.[173] Det tas derfor til orde for at det iverksettes en proaktiv, snarere enn en reaktiv, tilnærming til utfordringene. Eksemplene nedenfor viser at Norges erfaring med digitale verktøy og digital informasjon er ikke bare positiv.

I CrashCube-saken ble dataverktøyet «CrashCube» benyttet for å hente ut ulike typer data fra kjøretøyer. CrashCube kan etter utviklerens egen hjemmeside hente ut informasjon som kjøretøyhurtighet, setebelt bruk, tidsstempler, motorinformasjon, etc.[174] 20. desember 2019 sendte Riksadvokaten ut brev til politimesterne og Statens vegvesen om at det var oppdaget feilinformasjon i rapportene som senere ble brukt i straffesaker. Feilene gjeldt både uthenting og fortolkning av informasjon.[175] Dataverktøyet som sakkyndige i vegvesenet hadde benyttet presenterte med andre ord gale «out-put»-data som deretter ble lagt fram for en dømmende rett. Det ble naturlig nok iverksatt tiltak for å undersøke hvorvidt norske borgere kan ha blitt feilaktig dømt på bakgrunn av informasjonen. Undersøkelsene pågår ennå i skrivende stund.

I en annen sak fra 2019 ble det avdekket at teledata fra Danmark har vært brukt i norske straffesaker og at dataen fra teleleverandørene inneholdt feil som burde vært avslørt. Informasjon om personer og hvor de har befunnet seg i Danmark har blitt brukt i norske strafferettssaker og det er iverksatt undersøkelser for å prøve å avdekke hvorvidt

---

[172] **Utkast**: Konvensjonsforslag «Draft Convention on Electronic Evidence». Tilgjengelig på: https://www.researchgate.net/publication/309878298_Draft_Convention_on_Electronic_Evidence
[173] Eivind Koflaath og Magne Strandberg, Om utviklingen av robusthetskravet på side 45 i Hedlund, M. mfl. 1. utg. 2015. *Bevis i straffesaker: utvalgte emner*.
[174] Digipols hjemmeside. Om CrashCube varen. Tilgjengelig på: https://digitpol.com/crashcube/
[175] Riksadvokaten (20.12.19) *Statens vegvesen og deres bruk av dataverktøy ved sakkyndig bistand i straffesaker.* Riksadvokaten (20.12.19) *Statens vegvesen og deres bruk av dataverktøyet CrashCube.*

feilinformasjonen kan ha hatt innvirkning på resultatet i enkelte saker. Feilene kan strekke seg så langt tilbake som til 2010.[176] I begge sakene – CrashCube og teledata-saken – uttales det fra Riksadvokaten at det ikke eksisterer noen enkel måte å undersøke hvilke straffesaker feilinformasjonen har vært brukt i, eksempelvis gjennom automatiserte søk.[177]

Slik jeg ser det har politiet i utgangspunktet behov for å benytte ulike KI-systemer dersom de skal holde tritt med et kriminalitetsbilde i stadig utvikling. Samtidig er det et behov for kontroll med slike KI-systemer dersom de skal benyttes i bevisbehandling i straffesaker, jf. EU-kommisjonen.[178] Hensynet til effektiv kriminalitetsbekjempelse trekker således i en retning, mens hensynet til den tiltaltes rettssikkerhet trekker i den andre. Hensynet til tillit trekker imidlertid i begge retninger; dersom politiet møter barrierer som hindrer dem fra å bekjempe kriminaliteten synker tillitens til politiet, mens dersom rettssikkerhetshensynet neglisjeres synker tilliten til rettsvesenet.

På bakgrunn analysen i del 1 og rettskildene som er presentert i del 2, tas det til orde for opprettelsen av en rettsdigital kommisjon.

4.3 Den rettsdigitale kommisjonen

Det ble i forrige punkt pekt på hvorfor en rettsdigital kommisjon er nødvendig, og kanskje viktigst, hvorfor den kommer til å bli nødvendig i nær fremtid. Det som vil gjennomgås i dette punktet er hvilke oppgaver det er nødvendig at en slik kommisjon tildeles.

Som nevnt tas det utgangspunkt i hovedproblemet med KI-systemer i bevisbehandling, nemlig usikkerheten. Det er påvist et behov for kontroll for «high-risk» KI-verktøy samtidig som det antydes at det vil stilles fremtidige krav fra EU. Kravene er nødvendige fordi de rokker ved fundamentale rettigheter samtidig som bruken av slike systemer er ventet å vokse. I EU kommisjonens white paper adresseres det at statene må regne med å sørge for *«[r]equirements ensuring that the AI systems are robust and accurate, or at least correctly reflect their level of accuracy, during all life cycle phases.»*[179]

---

[176] Riksadvokaten (15.01.20) *Trafikkdata innhentet fra Danmark - videre oppfølging*.
[177] Op.cit. Øverst på side 2. «Det er et beskjedent antall saker som så langt er identifisert i Norge. Det er heller ikke registre over saker hvor det er innhentet trafikkdata fra utlandet, ei heller automatiserte søkemuligheter som med sikkerhet vil kunne fange opp alle saker hvor det er bedt om bistand fra utenlandske myndigheter.»
[178] EU kommisjonens *White Paper «on Artificial Intelligence: a European approach to excellence and trust.»* Side 17.
[179] *White Paper on Artificial Intelligence: a European approach to excellence and trust*. Side 20 bokstav d første dott.

En av hovedoppgavene til en rettsdigital kommisjon bør være testing og gjennomgang av de digitale verktøy som benyttes i straffeprosessen. Testing er en kontinuerlig prosess som krever mye tid og ressurser. Testingen krever samtidig store mengder uavhengig data som kan benyttes som testdata. Opprettelsen av en statlig og uavhengig kommisjon vil gjøre det lettere å tilby testdata som er vernet av personvernregler. Den digitale kommisjonen kan i kraft av å ha taushetsplikt utad få tilgang til den nødvendige dataen for testing av programvarene. På den måten kan programmer fra leverandører som ikke deler informasjon om programmet testes mer effektivt. I tilfeller der koder og treningsdata er tilgjengelig, eller blir gjort tilgjengelig for kommisjonen, kan kommisjonen gjennomgå programmet. Hvor den uavhengige kommisjonen helt konkret burde plasseres har jeg ingen spesiell formening om.

En rettsdigital kommisjon kan – i likhet med den rettsmedisinske – føre kontroll med sakkyndiguttalelser som legges fram for retten. På en mer generell basis kan kommisjonen stå for sakkyndigutdanning. [180] For å levere sakkyndiguttalelser kreves det at de sakkyndige har en inngående forståelse av sitt mandat og hva de skal utrede.

En annen oppgave kan være å utarbeide standarder for bruk av KI-systemene. Eksempelvis foreligger det i dag standarder på bruk av måleverktøy som benyttes ved fartskontroller av kjøretøyer, men ikke bruk av KI-verktøy.[181] I den forbindelse kan kommisjonen – i likhet med NRGD – bestå av eksperter og personer med juridisk utdanning. På den måten vil det kunne gis ekspertråd ved utarbeidelser av standarder, samtidig som det føres kontroll med at de etterleves. Som tidligere nevnt burde kommisjonens oppgave heller ikke begrense seg til kun KI-verktøy da det er kjent at det følger potensielle feilkilder ved nær sagt alle digitale etterforskningsverktøy.[182] Når det allerede er tatt til orde for at det kreves strenge former for kontroll ved alminnelige digitale etterforskningsverktøy kommer KI-problematikken på toppen av denne.

En annen oppgave vil være å bidra med ekspertise i diskusjonen rundt bruken av KI-systemer i straffeprosessen og utenfor. For eksempel arbeider NRGD med å ekspandere til sivil- og administrativ rett. På lengre sikt kan kommisjonen bidra med kunnskap til fordel for andre

---

[180] Eksempelvis har den rettsmedisinske kommisjonen som oppgave å utdanne sakkyndige etter Forskrift om Den rettsmedisinske kommisjon § 3 c.
[181] Politidirektoratet, 1 Instrukser for POLITIETS TRAFIKKTJENESTE GP-4027 Januar 2016. På side 21. Punkt 02.5 «instruks for bruk av laser ved fartsmåling».
[182] Erlandsen, T-E. *Fallacies when Evaluating Digital Evidence Among Prosecutors in the Norwegian Police Service.*
Borhaug, T. S. (2019). *The Paradox of Automation in Digital Forensics.*
Ekfeldt, J. (2016). *Om informationstekniskt bevis.*

departementer, ikke bare justis- og beredskapsdepartementet. Eksempelvis vil robotiseringen av skatteetaten kunne kontrolleres, i likhet med andre felter der maskinlæringsmodeller implementeres for administrative formål.

En problematikk som har en litt annen innfallsvinkel, men som kommer til å bli en utfordring på lik linje med KI-systemer som analyseverktøy, er utnyttelse av «deepfakes» teknikker. «Deepfake» er bevis (lyd, bilde, video, etc) som har blitt manipulert gjennom maskinlæringssystemer. Bevisene framstår som ordinære bevis og kan skape utfordringer for rettsvesenet i fremtiden etterhvert som teknologien forbedrer seg. Isabelle Ringnes som leder TENK[183] forklarer «deepfakes» slik:

«Deepfakes er videoer, bilder og lydklipp som har blitt manipulert digitalt, ved hjelp av maskinlæring. Slik kan man få det til å virke som om en hvilken som helst person har sagt noe vedkommende aldri ville sagt.»[184] Hun forklarer videre at *«[m]ot 100 personer som utvikler deepfakes, står det én person som jobber for å identifisere om en video er falsk eller ekte.»*

Dermed kan for eksempel seerne av et videoklipp tro at personen som imiteres ytrer noe absurd eller i den retningen manipulatorene ønsker. Eksempler finnes allerede i Norge, noe Dagbladet lagde en sak om tidligere i år. Se artikkelen «Denne mannen eksisterer ikke» i fotnote.[185] Teknikken fikk også mange til å åpne øynene da et klipp av Barack Obama gikk viralt, som i realiteten viste seg å være komikeren Jordan Peele.[186] Det er rimelig å anta at etterhvert som teknologien utvikles vil en før eller siden stå overfor slike bevis i rettssaler verden over.

Problemet med «deepfakes» er anerkjent i Norges «National strategi for KI» under kapitlet «kunstig intelligens i kriminalitetsbekjempelse» punkt 3. uttrykt som «... mulighet for å plante/forvrenge spor.» Vi er inne i en æra der digitalisering er den nye normalen for alt vi foretar oss. Det må dermed forventes at når teknologien tillater det juks, vil noen alltid prøve

---

[183] «Tech-nettverket for kvinner»
[184] Artikkel på online.no *«Derfor er deepfake-videoer en trussel mot demokratiet».* Tilgjengelig på https://www.online.no/sikkerhet/deep-fake-bekymrer-ekspertene?cid=p-prog_fix_Apollo_sikker_df1_aller (Sist lest 22.05.2020)

[185] Artikkel på dagbladet.no *«Denne personen eksiterer ikke».* Tilgjengelig på https://www.dagbladet.no/nyheter/denne-personen-eksisterer-ikke/72158593 (Sist lest 22.05.2020)

[186] Artikkel på theverge.com *«Watch Jordan Peele use AI to make Barack Obama deliver a PSA about fake news.»* Tilgjenglig på: https://www.theverge.com/tldr/2018/4/17/17247334/ai-fake-news-video-barack-obama-jordan-peele-buzzfeed (sist lest 29.05.20)

å jukse. Eksempelvis kan lyd- eller videoklipp brukes til å dra ned troverdigheten til fornærmede ved å få personen til å si eller gjøre noe som setter personen i disfavør. Dersom det erkjennes at dette før eller siden kommer for domstolene i form av bevis vil spørsmålet bli om Norge har mekanismer som kan oppdage dette. «Deepfakes» vil i fremtiden kunne true hensynet til et materielt riktig resultat, rettferdighet og tilliten til rettssystemet dersom problemet ikke kan adresseres. En rettsdigital kommisjon vil i den forbindelse kunne utvikle kompetansen til å avsløre forsøk på juks.

Listen av oppgaver den rettsdigitale kommisjonen kan tildeles er ikke uttømmende, men snarere noen av de mest presserende for dagens og morgendagens situasjon. Løsningen som foreslås er ikke enkel og vil koste penger. Imidlertid må kostandene sees i et nettoperspektiv der kostnadene som spares – ved å holde problematikken utenfor den enkelte sak – holdes opp mot kostnadene en rettsdigital kommisjon vil føre med seg. På bakgrunn av Riksadvokatens oppfordring til å automatisere og effektivisere politiets arbeid gjennom digitale løsninger virker det som viljen til å modernisere politiet er stor.[187] På samme tid virker det som baksiden av medaljen blitt oversett i det offentlige rom. Ved å automatisere kritiske deler av straffeprosessen vil man også måtte forholde seg til paradokset ved automatisering: *«the more efficient the automated system, the more crucial the human contribution of the operators. Humans are less involved, but their involvement becomes more critical"*.[188]

Dette er noe norsk straffeprosess må forholde seg til dersom automatiseringen av kritiske funksjoner skal fortsette. Det er derfor viktig at ansvaret ikke neglisjeres og pulveriseres av aktørene i strafferettspleien.

[187] Riksadvokatens notat 07.01.2019. *Notat om utviklingen ved etterforskningsfelt.* Punkt 4.3.1. Side 9.
[188] Borhaug, T. S. (2019). *The Paradox of Automation in Digital Forensics.* (Masteroppgave, NTNU, Norges teknisk-naturvitenskapelige universitet). Side 22. Tilgjengelig på: https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2617753 Definisjonen er utarbeidet av Josh Kaufman.

5 Kildeliste

5.1 Lov- og forskriftsregister

Lover:

Grl. Lov 17 mai 1814 Kongeriket Norges Grunnlov. (Grunnloven).

Vtrl. Lov 18 juni 1965 nr. 4 om vegtrafikk (Vegtrafikkloven).

Strpl. Lov 22 mai 1981 nr. 25 om rettergangsmåten i straffesaker (Straffeprosessloven).

Politil. Lov 4 august 1995 nr. 53 om politiet (Politiloven).

Mrl. Lov 21 mai 1999 nr. 30 om styrking av menneskerettighetenes stilling i norsk rett (Menneskerettloven).

Strl. Lov 20 mai 2005 nr. 28 om straff (Straffeloven).

Lov 15 juni 2018 nr. 38 om behandling av personopplysninger (Personopplysningloven).

Forskrifter:

Forskrift 28. juni 1985 nr. 1679 om ordningen av påtalemyndigheten (Påtaleinstruksen).

Forskrift 13. februar 2018 nr. 240 om Den rettsmedisinske kommisjon.

5.2 Forarbeider og andre offentlige dokumenter

Meld. St. 28 (2018-2019) Datatilsynets og Personvernsnemdas årsrapporter for 2018.

NUT 1969: 3 Innstilling om rettergangsmåten i straffesaker fra Straffeprosesslovkomiteen (Komiteen til revisjon av straffeprosessloven).

NOU 2001: 12 Rettsmedisinsk sakkyndighet i straffesaker.

NOU 2016: 24 Ny straffeprosesslov.

Ot.prp. nr. 90 (2003-2004) Om lov om straff (straffeloven)

Politidirektoratet, 1 Instrukser for POLITIETS TRAFIKKTJENESTE GP-4027 Januar 2016.

Regjeringen (14.01.20), Publikasjonskode: H-2458 B, *Nasjonal strategi for kunstig intelligens*, Kommunal- og moderniseringsdepartementet, Oslo.
Tilgjengelig på: regjeringen.no

Riksadvokatens rundskriv nr. 3/2018. (Revidert 21.02.19) *Kvalitetskrav til straffesaksbehandlingen i politiet og ved statsadvokatembete mv. (kvalitetsrundskrivet).* Tilgjengelig på*:* https://www.riksadvokaten.no/document/nytt-kvalitetsrundskriv/

Riksadvokatens notat (07.01.19) *Notat om utviklingen ved etterforskningsfelt.* Tilgjengelig på: https://www.riksadvokaten.no/document/utvikling-av-etterforskingsfeltet/

Riksadvokaten (20.12.19) *Statens vegvesen og deres bruk av dataverktøy ved sakkyndig bistand i straffesaker.* Tilgjengelig på:
https://www.riksadvokaten.no/document/vegvesenet-mulig-feil-i-trafikkdataverktoy/

Riksadvokaten (20.12.19) *Statens vegvesen og deres bruk av dataverktøyet CrashCube.* Tilgjengelig på*:* https://www.riksadvokaten.no/document/vegvesenet-mulig-feil-i-trafikkdataverktoy/

Riksadvokaten (15.01.20) *Trafikkdata innhentet fra Danmark - videre oppfølging*.
Tilgjengelig på: https://www.riksadvokaten.no/document/mulig-feilinformasjon-i-trafikkdata-fra-danmark-videre-oppfolging/

5.3 Rettspraksis

Høyesterettspraksis:

Rt. 1982 s. 1264

Rt. 1983 s. 1275

Rt. 1984 s. 840

Rt. 1985 s. 168

Rt. 1990 s. 1008

Rt. 1994 s. 610

Rt. 1996 s. 1114

Rt. 2001 s. 543

Rt. 2002 s. 1744

Rt. 2003 s. 1682

Rt. 2005 s. 1353

Rt. 2006 s. 856

Rt. 2007 s. 10

Rt. 2007 s. 1255

Rt. 2008 s. 605

Rt. 2010 s. 655

HR-2011-1969-U

Rt. 2012 s. 897

Rt. 2013 s. 905

5.4 Internasjonale konvensjoner og utenlandsk lov

Konvensjoner:

Convention for the Protection of Human Rights and Fundamental Freedoms, Rome, 04.11.1950. (Menneskerettskonvensjonen)

UNs International Covenant on Civil and Political rights, 16.12.1966.

Convention on Cybercrime, Budapest, 23.11.2001. Tilgjengelig på: https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/185

Utkast: Konvensjonsforslag «Draft Convention on Electronic Evidence». Tilgjengelig på:

https://www.researchgate.net/publication/309878298_Draft_Convention_on_Electronic_Evidence

Amerikansk lov:

U.S Federal Rules of Evidence (FRE)", Rule 702. «Testimony by Expert Witnesses». Tilgjengelig på: https://www.law.cornell.edu/rules/fre/rule_702

5.4.1 Internasjonal rettspraksis

Den europeiske menneskerettsdomstols praksis (EMD):

Van Mechelen and others v. The Nederlands, 30 October 1997, (Saksnummer 55/1996/674/861-864).

Jasper v. United Kingdom, Grand Chamber 16 February 2000, (Saksnummer 27052/95).

Gül v. Turkey, 14 December 2000, (Saksnummer 22676/93).

Al Khawaja and Tahery v. United Kingdom, Grand Chamber 15. Desember 2011, (Saksnummer 26766/05 og 22228/06).

Sigurður Einarsson and others v. Iceland, 4 September 2019, (Saksnummer 39757/15).

Amerikansk rettspraksis:

Eric Loomis vs State of Wisconsin, cert. denied, 137 S.Ct. 2290 (2017). Wisconsin Supreme Court. Appeal to the United States Supreme Court denied. Tilgjengelig på: https://www.courts.ca.gov/documents/BTB24-2L-3.pdf

5.5 Bøker

Andenæs, J. og Myhrer, T. 4. utg. 2009. *Norsk straffeprosess*. Oslo: Universitetsforlaget.

Bendiksen, C. og Norman Hansen, E. 1. utg. 2019. *Når juss møter AI.* Oslo: Gyldendal.

Hedlund, M. Mfl. 1. utg. 2015. *Bevis i straffesaker: utvalgte emner.* Oslo: Gyldendal juridisk.

Kolflaath, E. 1. utg. 2013. *Bevisbedømmelse i praksis.* Bergen: Fagbokforlaget Vigmostad & Bjerke AS.

Løvlie, A. 1. utg. 2014. *Rettslige faktabegreper.* Oslo: Gyldendal Juridisk.

Mitchell, T. 1. utg. 1997. *Machine Learning*. McGraw-Hill Science/Engineering/Math.

Russel, S og Norvig, P. 3. utg. 2010. *Artificial Intelligence: A Modern Approach.* New Jersey: Prentice Hall.

Saltzer, J. and Frans Kaashoek. 1. utg. 2009. *Principles of Computer System Design: An Introduction.* Burlington: Morgan Kaufmann Publisher.

Øyen, Ø. 2. utg. 2016. *Straffeprosess.* Bergen: Fagbokforlaget.

Årnes, A. mfl. 1. utg. 2018. *Digital Forensics*. Hoboken: John Wiley & Sons Ltd.

5.6 Avhandlinger

Borhaug, T. S. (2019). *The Paradox of Automation in Digital Forensics.* (Masteroppgave, NTNU, Norges teknisk-naturvitenskapelige universitet). Tilgjengelig på: https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2617753

Ekfeldt, J. (2016). *Om informationstekniskt bevis.* (Doktoravhandling, Stockholms universitet). Tilgjengelig på: http://su.diva-portal.org/smash/get/diva2:900594/FULLTEXT05.pdf

Erlandsen, T-E. (2019). *Fallacies when Evaluating Digital Evidence Among Prosecutors in the Norwegian Police Service.* (Masteroppgave, NTNU, Norges teknisk-naturvitenskapelige universitet). Tilgjengelig på: https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2617771

Strandberg, M, (2010) *Beviskrav i sivile saker. En bevisteoretisk studie av den norske beviskravslærens forutsetninger*. (Doktoravhandling, Universitetet i Bergen).

5.7 Rapporter

Datatilsynet (2018), *Kunstig intelligens og personvern.* Oslo. Tilgjengelig på: www.datatilsynet.no

Deloitte (2018), *Policing 4.0 Deciding the future of policing in the UK,* London: The Creative Studio at Deloitte.
Tilgjengelig på: www2.deloitte.com

House of Lords (2018), *AI in the UK: ready, willing and able?* London. Tilgjengelig på: https://publications.parliament.uk/

Interpol/UNICRI (2019), *artificial intelligence and robotics for law enforcement.* Tilgjengelig på: www.unici.it

5.8 Artikler

Angwin, J., Larson, J., Mattu, S. & Kirchner, L. (2016*). Machine Bias.* Tilgjengelig på: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing, (sist lest 28.05.20).

Bass, D. and Huet, E. (2017). *Researchers Combat Gender and Racial Bias in Artificial Intelligence.* Tilgjengelig på: https://www.bloomberg.com/news/articles/2017-12-04/researchers-combat-gender-and-racial-bias-in-artificial-intelligence, (sist lest 08.05.20).

Friedman, B. og Nissenbaum, H. (1996) *Bias in computer systems.* Publisert: ACM Transactions on Information Systems, juli 1996.

Horsman, G. (2019) *Tool testing and reliability issues in the field of digital forensics.* Publisert i: *Digital Investigation 01.03.2019. p. 163-175.*

Huber, J-A., Memminger, M., Soppitt, M. og Hayday, M. (2019). *Cutting Through Complexity In Financial Crimes Compliance.* Tilgjengelig på: https://www.forbes.com/sites/sap/2020/05/28/leadership-times-of-crisis-the-importance-of-empathy-and-innovation/#2bf93aae5857 (sist lest 21.05.20)

Meyer, B, McCarthy, J. (2011). *Communications of the ACM* (28. Oktober 2011): https://cacm.acm.org/blogs/blog-cacm/138907-john-mccarthy/fulltext, (sist lest 28.05.20).

Naarttijärvi, M. (2017). *Rättsstatlighet och algoritmiska svarta lådor.* I: Örjan Edström, Johan Lindholm & Ruth Mannelqvist (ed.), Jubileumsskrift till Juridiskainstitutionen 40 år (s. 245-259). Umeå: Juridiska institutionen, Umeå universitet.

Rui, J. P. (2014). *Straffeprosessen i perspektiv.* Publisert i: *Jussens venner. Juni 2014 (Volum 49).*

Saks, M. J. og Koehler, J. J. (2005) *The Coming Paradigm Shift in Forensic Identification Science.* Publisert i *Science,* 5. august 2005.

Searston, R. A. og Chin, J. (2019) *The Legal and Scientific Challenge of Black Box Expertise.* Publisert i: *University of Queensland Law Journal.* The University of Sydney Law School.

Shalaginov, A. Og Franke, K. (2017) *A deep neuro-fuzzy method for multi-label malware classification and fuzzy rules extraction.* Publisert i: 2017 IEEE Symposium Series on Computational Intelligence (SSCI).

Vestby, A. og Vestby, J. (2019). J. *Machine Learning and the Police: Asking the Right Questions.* Publisert i: *Policing: A Journal of Policy and Practice.* Juni 2019.

5.9 EU publikasjoner

EU Commission (08.04.2019). *Ethics Guidelines for Trustworthy AI, High-Level Expert Group on Artificial Intelligence.* Brussel.

EU Commission (19.02.2020) COM (2020) 65 final. *White Paper on Artificial Intelligence: a European approach to excellence and trust*, Brussel.

The European Parliament and of the Council. Directive 2016/680 (27.04.2016). *On the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free*

movement of such data, and repealing Council Framework Decision 2008/977/JHA. (Law Enforcement Directive)

The European Parliament and of the Council. Regulation, (Date made: 27.04.2016), (Implementation date: 25.05.2018). COM/2012/010 final – 2012/0010 (COD), *General Data Protection Regulation (GDPR).* Forordningen er inntatt i lov 15 juni 2018 nr. 38 om behandling av personopplysninger (personopplysningloven) § 1.

5.10 Øvrige kilder

5.10.1 Nettsteder

Artikkel på politiforum.no. Intervju med Katrin Franke, professor ved NTNU. *«Palantir holder planene for framtiden hemmelige. Professor er bekymret for at data kan havne i private selskapers hender.»* Tilgjengelig på: https://www.politiforum.no/artikler/palantir-holder-planene-for-framtiden-hemmelige-professor-er-bekymret-for-at-data-kan-havne-i-private-selskapers-hender/434026 (sist sett 21.05.20).

Automatisk skiltgjenkjenning (ANPR) Tilgjengelig på: https://www.vegvesen.no/fag/fokusomrader/trafikksikkerhet/skiltleser (sist lest 29.05.20).

Artikkel på politiforum.no *«Streng kontroll med brukerne av politiets nye analyseverktøy»* Tilgjengelig på: https://www.politiforum.no/artikler/streng-kontroll-med-brukerne-av-politiets-nye-analyseverktoy/404492 (sist lest 21.05.20).

Artikkel på politiforum.no *«Norsk forskning på framtidens politi til topps i Interpol»* Tilgjengelig på: https://www.politiforum.no/artikler/norsk-forskning-pa-framtidens-politi-til-topps-i-interpol/449053 (sist lest 21.05.20).

Store norske leksikon (www.snl.no)

The Netherlands Register of Court Experts (NRGD). Hjemmeside: https://english.nrgd.nl/

Digitpols hjemmeside. Utviklerne av dataverktøyet CrashCube .https://digitpol.com/crashcube/

Artikkel på online.no *«Derfor er deepfake-videoer en trussel mot demokratiet»*. Tilgjengelig på: https://www.online.no/sikkerhet/deep-fake-bekymrer-ekspertene?cid=p-prog_fix_Apollo_sikker_df1_aller (sist lest 22.05.20).

Artikkel på dagbladet.no *«Denne personen eksiterer ikke».* Tilgjengelig på: https://www.dagbladet.no/nyheter/denne-personen-eksisterer-ikke/72158593 (sist lest 22.05.20)

Aftenpostens artikkel: *«Oslo-politiet tester utrykning til oppdrag med droner fra fredag»* Tilgjengelig på: https://www.aftenposten.no/osloby/i/LAwEzp/oslo-politiet-tester-utrykning-til-oppdrag-med-droner-fra-fredag (sist lest 29.05.20)

Artikkel fra theverge: «*Watch Jordan Peele use AI to make Barack Obama deliver a PSA about fake news»* Tilgjengelig på: https://www.theverge.com/tldr/2018/4/17/17247334/ai-fake-news-video-barack-obama-jordan-peele-buzzfeed (sist lest 29.05.20)

5.11 Personlige meddelelser

Personlige meddelelser, Katrin Franke, professor i computer science ved NTNU (Norges teknisk-naturvitenskapelige universitet). 2019 og 2020. Om eksempler angående kunstig intelligens i dataetterforskning. [Gjengitt med samtykke].